

AUS DEM LEHRSTUHL  
FÜR HUMANGENETIK  
PROF. DR. BERNHARD WEBER  
DER FAKULTÄT FÜR MEDIZIN  
DER UNIVERSITÄT REGENSBURG

# SUCHE NACH EINEM NEUEN BRUSTKREBSGEN MITTELS POSITIONELLER KANDIDATENGENANALYSE

Inaugural-Dissertation  
zur Erlangung des Doktorgrades  
der Medizin

der  
Fakultät für Medizin  
der Universität Regensburg

vorgelegt von  
Monika Huber

2011



AUS DEM LEHRSTUHL  
FÜR HUMANGENETIK  
PROF. DR. BERNHARD WEBER  
DER FAKULTÄT FÜR MEDIZIN  
DER UNIVERSITÄT REGENSBURG

# SUCHE NACH EINEM NEUEN BRUSTKREBSGEN MITTELS POSITIONELLER KANDIDATENGENANALYSE

Inaugural-Dissertation  
zur Erlangung des Doktorgrades  
der Medizin

der  
Fakultät für Medizin  
der Universität Regensburg

vorgelegt von  
Monika Huber

2011

Dekan: Prof. Dr. Dr. Torsten E. Reichert

1. Prüfer: Prof. Dr. Bernhard Weber

2. Prüfer: Prof. Dr. Olaf Ortmann

Tag der mündlichen Prüfung: 01.12.2011

FÜR MEINE ELTERN

## Inhaltsverzeichnis

<b>1</b>	<b>EINLEITUNG</b>	<b>4</b>
1.1	Karzinogenese . . . . .	4
1.2	Grundlagen des familiären Mammakarzinoms . . . . .	6
1.2.1	Epidemiologie des Brustkrebses . . . . .	6
1.2.2	Risikofaktoren für Brustkrebs . . . . .	7
1.2.3	Erblichkeit und Familiarität . . . . .	7
1.3	Strategien zur Identifizierung von Krankheitsgenen . . . . .	13
1.4	Fragestellung der vorliegenden Arbeit . . . . .	14
<b>2</b>	<b>MATERIAL, PATIENTEN UND METHODEN</b>	<b>16</b>
2.1	Verwendete Geräte, Software und Reagenzien . . . . .	16
2.1.1	Geräte und Software . . . . .	16
2.1.2	Reagenzien . . . . .	17
2.2	Patienten . . . . .	18
2.3	Die Methode der familiären Kopplungsanalyse . . . . .	19
2.3.1	Die Kopplungsanalyse . . . . .	19
2.3.2	Kopplungsanalyse mit 50K XbaI-GenChip . . . . .	22
2.3.3	Bestätigung und weitere Eingrenzung der Kopplungsregionen mit Mikrosatelliten-Markern . . . . .	25
2.4	Analyse der Kandidatengene . . . . .	32
2.5	Sequenzierung von Kandidatengenomen . . . . .	32
2.5.1	Das Prinzip der DNA-Sequenzierung . . . . .	32
2.5.2	Durchführung der DNA-Sequenzierung . . . . .	34
<b>3</b>	<b>ERGEBNISSE</b>	<b>38</b>
3.1	Familiäre Kopplungsanalyse . . . . .	38
3.1.1	Annahme 1: Alle an Brustkrebs erkrankte Familienmitglieder tra- gen den genetischen Defekt . . . . .	38
3.1.2	Annahme 2: Eine der betroffenen Personen im Stammbaum er- krankte an einem sporadischen Brustkrebs . . . . .	40
3.1.3	Zusammenstellung der Kandidatenregionen . . . . .	51
3.2	Kandidatengene . . . . .	51
3.3	Sequenzierung . . . . .	54
<b>4</b>	<b>DISKUSSION</b>	<b>55</b>
4.1	Bilanz und Ausblick . . . . .	65
<b>5</b>	<b>ZUSAMMENFASSUNG</b>	<b>67</b>
<b>6</b>	<b>LITERATURVERZEICHNIS</b>	<b>68</b>
<b>7</b>	<b>ABBILDUNGSVERZEICHNIS</b>	<b>75</b>
<b>8</b>	<b>ANHANG</b>	<b>76</b>
8.1	Abkürzungen . . . . .	76
8.2	Auswertungen . . . . .	79
<b>9</b>	<b>ERKLÄRUNGEN</b>	<b>117</b>

---

9.1	Ehrenwörtliche Erklärung . . . . .	117
9.2	Erklärung zur Anregung und Überwachung der Dissertation . . . . .	117
<b>10</b>	<b>DANKSAGUNG</b>	<b>118</b>
<b>11</b>	<b>LEBENS LAUF</b>	<b>119</b>

# 1 EINLEITUNG

## 1.1 Karzinogenese

Ein Tumor entsteht durch die Vermehrung von körpereigenen entarteten Zellen. Ein allgemeines Kennzeichen fast aller Krebszellen ist eine genomische Instabilität, verursacht entweder durch vererbte Mutationen in Genen, die die genomische Integrität überwachen, oder durch Mutationen, die sich in somatischen Zellen während der Tumorentstehung ansammeln [1]. Die Ursachen von Krebs sind meist komplexe Aberrationen im Genom der entarteten Zellen. Solche Aberrationen erstrecken sich auf Mutationen in kodierenden und regulatorischen Sequenzen, Veränderungen der Ploidie, kleine genomische Vervielfältigungen, große Amplifikationen, strukturelle Umordnungen, homozygote Deletionen und Verlust der Heterozygotität. Die durch die genetischen Veränderungen ausgelöste Transformation von einer normalen Zelle zu einer Tumorzelle geht mit einer Störung der zellulären Regulations- und Abwehrmechanismen einher [2]. Krebs ist somit das Ergebnis einer Störung unter anderem der genomischen Integrität, der Wachstumskontrolle, der zellulären Differenzierung, der Proliferation und der Apoptose.

Krebs entsteht als Ergebnis einer Abfolge von somatischen Mutationen, teilweise unter Beteiligung einer vererbten Prädisposition [3]. Nach der Mehrschritt-Theorie der Krebsentstehung entwickelt sich ein maligner Tumor meist über viele Jahre durch Ansammlung einer Reihe von Defekten in verschiedenen Genen. Eine Mutation verleiht der Zelle, in der sie sich ereignet, einen klonalen Selektionsvorteil und stellt eine erste Stufe in Richtung Malignität dar. Die Akkumulation einer spezifischen Kombination genetischer Veränderungen bis zu einem kritischen Punkt verursacht schließlich den malignen Phänotyp. Nach der klonalen Entwicklungstheorie geht das Tumorwachstum in der Regel von einer einzigen transformierten somatischen Zelle aus. Die Zelle proliferiert und bildet eine zunächst monoklonale Zellformation. Die genetische Instabilität der Tumorzellen führt jedoch schließlich zu einer zunehmenden Tumorerheterogenität [2]. Durch zusätzliche Mutationen in den klonalen Tumorzellen entstehen Subklone, die unterschiedliche Fähigkeiten zur Infiltration und Metastasierung aufweisen.

Eine genetische Prädisposition ist bei 5-10% aller menschlichen Tumore verantwortlich bzw. mitverantwortlich. Familiäre Tumorsyndrome werden durch eine Keimbahnmutation in einem Tumorgen verursacht. Die Mehrheit der vererbten Krebsgene wird scheinbar in den meisten adulten Geweben exprimiert, obwohl sich bei Trägern einer Keimbahnmutation in einem solchen Gen oft nur ein begrenztes Spektrum an Krebs manifestiert [4]. Obwohl das Konzept von verzweigten Signalwegen oder alternativer Genfunktion in unterschiedlichen Geweben hilft, die Gewebespezifität des Krebspektrums in Mutationsträgern zu erklären, scheint die Erklärung noch unvollständig. Die Phänotypen zwischen Individuen, die dieselbe Keimbahnmutation tragen, variieren. Dies ist zu beobachten, da andere genetische und Umweltfaktoren zusätzlichen Einfluss ausüben.

Als klassische Tumorgene werden Proto-onkogene/Onkogene und Tumorsuppressorgene verstanden. Darüber hinaus können auch DNA-Reparaturgene, Apoptosegene sowie Telomerasegene zu Tumorgenen transformieren.

### Proto-onkogene/Onkogene

Proto-onkogene sind normale zelluläre Gene, deren Genprodukte Proliferation, Mobilität und Differenzierung von Zellen regulieren [2]. Sie sind meist eine Komponente eines komplexen Netzwerks der intrazellulären Signaltransduktion, an dessen Ende die mitotische



Teilung der Zelle steht. Proto-onkogene werden nach der Funktion ihrer Genprodukte in fünf Gruppen eingeteilt [3]:

- sezernierte Wachstumsfaktoren
- Wachstumsfaktor-Rezeptoren und Rezeptoren an der Zelloberfläche
- Komponenten der intrazellulären Signaltransduktion
- nukleäre DNA-bindende Proteine und Transkriptionsfaktoren
- Komponenten im Netzwerk der Zykline, zyklinabhängige Kinasen und Kinaseinhibitoren.

Es gibt verschiedene Mechanismen, die Proto-onkogene zu einem Onkogen aktivieren können. Amplifikation, Punktmutation, chromosomale Umstrukturierung, die zu einem neuen chimären Gen führt, Translokation in einen Bereich mit aktiver Transkription und Deregulierung der Genexpression sind mögliche genetische Veränderungen. Das Ergebnis ist eine deregulierte, gesteigerte Funktion (*gain of function*) des Onkogens bzw. Onkoproteins. Der Defekt wirkt sich dominant aus, d.h. ein einziges mutiertes Allel beeinträchtigt den Phänotyp der Zelle. Zurzeit sind über 100 Onkogene bekannt, die bei Tumoren durch Mutationen dominant aktiviert werden [3].

### **Tumorsuppressorgene**

Tumorsuppressorgene sind normale zelluläre Gene, deren Genprodukte negative Regulatoren z.B. von Wachstum und Zellzyklus sind [2]. Eine fehlende Bildung des Genprodukts bzw. ein fehlerhaftes Genprodukt bewirkt den Funktionsverlust (*loss of function*) des Tumorsuppressors und damit ein dereguliertes Wachstum der Zellen. Genetische Veränderungen zeigen ein rezessives Verhalten, d.h. eine tumorpromovierende Wirkung entsteht durch Veränderungen beider Allele in einem Tumorsuppressorgen. Ein Modell dafür ist die zwei-Treffer (*two hit*)-Hypothese der biallelischen Geninaktivierung beim Retinoblastom [5, 6]:

Das Retinoblastom ist ein aggressiver Tumor im Kindesalter. Etwa 40% sind familiär bedingt und zeigen eine dominante Vererbung mit unvollständiger Penetranz. Das familiäre Retinoblastom tritt im Gegensatz zum sporadischen häufig beidseitig auf und ereignet sich in einem früheren Lebensalter. Knudson stellte 1971 [5] die Hypothese auf, dass alle Retinoblastome zwei Treffer erfordern, von denen bei der familiären Form einer bereits vererbt wird. Das familiäre Retinoblastom ist in diesem Modell das Ergebnis von zwei Mutationen, wobei der Defekt an einem Allel durch eine Keimbahnmutation vererbt wurde, die maligne Transformation jedoch erst durch eine somatische Mutation des zweiten Allels ausgelöst wird. Die Funktion des später identifizierten Tumorsuppressors, das Retinoblastomgen (Rb1), liegt in der Kontrolle des Zellzyklus in der G1/S-Phase. Es kann jedoch nicht immer darauf geschlossen werden, dass familiäre und sporadische Krebsformen dem gleichen Entwicklungsprozess folgen.

Tumorsuppressorgene können durch Punktmutationen oder Deletionen (Verlust der Heterozygotie) inaktiviert werden. Ein weiterer Mechanismus ist die Methylierung des Promotors, die die Transkription des Gens stilllegt [3].

### **DNA-Reparaturgene [2, 3]**

Die DNA-Reparatur dient der Behebung von Schäden der DNA, die während der DNA-Replikation oder als Folge mutagener Einflüsse auftreten können. Reparatursysteme sind

in der Lage, DNA-Schäden zu erkennen und durch die reguläre Sequenz zu ersetzen. Während des Reparaturprozesses wird der Zellzyklus solange unterbrochen bis der Schaden behoben ist oder der programmierte Zelltod wird eingeleitet, wenn sich der Schaden nicht reparieren lässt. Defekte in Genen des Reparatursystems führen zu einer Anhäufung von Mutationen und genetischer Instabilität. Es gibt unter anderem Defekte der Nukleotidexzisionsreparatur, der Basenexzisionsreparatur, Defekte bei der Reparatur von DNA-Doppelstrangbrüchen und Defekte bei der Reparatur von Replikationsfehlern. Die erhöhte Mutationsrate erleichtert eine Tumorentstehung. DNA-Reparaturgene wirken in der Regel als Tumorsuppressoren.

### **Apoptosegene [2]**

Apoptose bezeichnet den Prozess des programmierten Zelltodes, der durch die betroffene Zelle selbst gesteuert und initiiert wird. Sie ermöglicht ein natürliches Gleichgewicht zwischen Zellvermehrung und Zellelimination. Apoptose kann durch eine Reihe von Zellschädigungen, durch spezifische Signale, bei Fehlen von Wachstumsfaktoren oder inadäquater Wachstumsstimulation ausgelöst werden. Veränderungen in Apoptosegenen führen zu einer Störung der Regulation der Apoptose oder der Apoptose selbst und sind mitverantwortlich für die Krebsentstehung.

### **Telomerasegene [2, 3]**

Die Teilungsfähigkeit menschlicher Zellen ist normalerweise zeitlich eingeschränkt, da sich bei jeder Zellteilung die Endabschnitte der Chromosomen – die Telomere – um etwa 50–100 bp verkürzen. Tumorzellen besitzen jedoch die Fähigkeit, sich unbegrenzt zu teilen; sie erlangen einen Zustand der Immortalisierung. Dieses Phänomen ist das Ergebnis einer Enzymaktivität (Telomerase), die verlorengegangene Telomerabschnitte ergänzen und stabilisieren kann. Neben Keimzellen können einige Tumorzellen die dafür verantwortliche Telomerase produzieren.

## **1.2 Grundlagen des familiären Mammakarzinoms**

### **1.2.1 Epidemiologie des Brustkrebses**

Brustkrebs ist in Deutschland wie auch weltweit die häufigste Karzinomdiagnose der Frau, die ein Lebenszeitrisko von 9,2% besitzt an Brustkrebs zu erkranken. Daraus lässt sich errechnen, dass jede elfte Frau in Deutschland im Verlauf ihres Lebens davon betroffen sein wird. Männer erkranken im Vergleich 50- bis 100-mal seltener an Brustkrebs [7].

Nach Schätzung der “Dachdokumentation Krebs” am Robert Koch-Institut wurde im Jahr 2004 bei über 57.000 Frauen die Diagnose Mammakarzinom gestellt, von denen etwa 23.200 unter 60 Jahre alt waren. Das entsprach einem Anteil von etwa 27,8% an allen Krebsneuerkrankungen bei Frauen. Die Inzidenz stieg in den letzten Jahren stetig an. Das mittlere Erkrankungsalter lag bei 63 Jahren und damit sechs Jahre unter dem durchschnittlichen Krebs-Erkrankungsalter für alle Krebsarten [8].

Über alle Brustkrebsstadien betrachtet, beträgt die relative 5-Jahres-Überlebensrate für Mammakarzinompatientinnen heute etwa 81% [8]. Die Prognose bei Brustkrebs ist somit gegenüber anderen bösartigen Neubildungen relativ gut. Seit Mitte der 1990er Jahre beobachtet man trotz eines Anstiegs der Inzidenz eine rückläufige Sterberate. Dieser Trend beruht unter anderem auf der Verbesserung der Diagnosetechnik, adäquaten Früherkennungsmaßnahmen und einem gestiegenen Bewusstsein für das Mammakarzinom in

der Bevölkerung. So wird ermöglicht, dass die Erkrankung frühzeitig aufgespürt werden kann und damit mit einer besseren Überlebensaussicht verbunden ist. Trotzdem liegt das Mammakarzinom auf Platz eins der Krebstodesursachenstatistik der Frauen und ist für 17,9% der Mortalität an Krebs verantwortlich. Im Jahre 2004 waren es 17.592 Frauen in Deutschland, die an Brustkrebs starben [8].

### 1.2.2 Risikofaktoren für Brustkrebs

Mammakarzinome treten am häufigsten postmenopausal auf und sind zum größten Teil sporadischer Genese, d.h. unabhängig von vererbten genetischen Faktoren. Für die Entstehung von Brustkrebs ist eine große Anzahl von Risikofaktoren bekannt. Dazu gehören allgemein höheres Lebensalter, frühe Menarche, späte Menopause, Nullipara oder ein höheres Alter bei der Geburt des ersten Kindes. Umgekehrt scheinen ausgetragene Schwangerschaften in jungen Jahren, mehrere Geburten und längere Stillzeiten das Brustkrebsrisiko zu verringern [8, 9]. Östrogen- und progesteronhaltige Ovulationshemmer nehmen nur einen geringfügigen Einfluss, während bei der postmenopausalen Hormonersatztherapie die Indikation genau geprüft werden sollte [10, 11]. Adipositas, Bewegungsarmut, Ernährung und ionisierende Strahlung (z.B. Mammographie-Screening) scheinen ebenfalls mit einer Risikoerhöhung verbunden zu sein.

Auf der anderen Seite haben epidemiologische Studien gezeigt, dass bei weiblichen Verwandten ersten Grades von Frauen, die an Brustkrebs erkrankt sind, das Risiko für bösartige Neubildungen in der Brust im Vergleich zur Allgemeinbevölkerung um etwa das Doppelte erhöht ist [12]. Die familiäre Häufung kann Ausdruck eines einzelnen Gendefekts, aber auch eines multifaktoriellen Geschehens unter Beteiligung exogener Faktoren und/oder mehrere Gene sein. Brustkrebs gehört wie die meisten Krebsarten zur Gruppe der komplexen Erkrankungen, in dem Sinne, dass er durch das Zusammenspiel von Genen und Umweltfaktoren entsteht [13].

### 1.2.3 Erbllichkeit und Familiarität

Eine genetische Disposition für Brustkrebs wird bei etwa 5-10% der erkrankten Frauen als Ursache gesehen [14]. In den letzten Jahren sind zahlreiche Untersuchungen durchgeführt worden, mit dem Ziel, die genetische Basis hinter dieser Prädisposition aufzuklären. Mittlerweile ist eine Einteilung der gefundenen Suszeptibilitätsallele für Brustkrebs in drei definierte Klassen verfügbar: seltene Allele mit hoher Penetranz, seltene Allele mit mäßiger Penetranz und verbreitete Allele mit niedriger Penetranz [12].

Folgende Tabelle stellt die Klassen und Hauptmerkmale der Suszeptibilitätsallele nach Stratton und Rahman (2008) [12] zusammen.

	Suszeptibilitätsgene mit hoher Penetranz	Suszeptibilitätsgene mit mäßiger Penetranz	Suszeptibilitätsgene mit niedriger Penetranz
Risikovarianten	multiple; verschiedene Mutationen	multiple; verschiedene Mutationen	SNPs
Frequenz	selten $\leq 0.1\%$	selten $\leq 0.6\%$	verbreitet 5-50%
Relatives Brustkrebsrisiko	10 - 20 fach	2 - 4 fach	bis zu 1,25-fach (heterozygot) oder 1,65-fach (homozygot)
Primäre Identifizierungs-Strategie	Kopplungsanalyse, Positionsklonierung	Direkte Untersuchung von Kandidatengenen nach kodierenden Variationen	Assoziationsstudie
Suszeptibilitätsgene bzw. -allele	BRCA1, BRCA2 TP53, PTEN STK11/LKB1, CDH1	ATM, CHEK2 BRIP1, PALB2	rs1045485 (CASP8) rs2981582 (FGFR2) rs3803662 (TOX3) rs889312 (MAP3K1) rs3817198 (LSP1) rs13281615 (8q) rs13387042 (2q)

### Seltene Suszeptibilitätsgene für Brustkrebs mit hoher Penetranz

Zu den Hochrisiko-Brustkrebs-suszeptibilitätsgenen zählen BRCA1, BRCA2, TP53, PTEN, STK11/LKB1 und CDH1.

#### ***BRCA1 und BRCA2***

Die beiden wichtigsten Brustkrebsgene BRCA1 und BRCA2 wurden in den 1990er Jahren durch Kopplungsanalyse und anschließende Positionsklonierung identifiziert [15–18]. BRCA1 ist auf Chromosom 17q21, BRCA2 auf Chromosom 13q12 lokalisiert. Pathogene Mutationen in BRCA1 und BRCA2 gehen mit einem etwa 10- bis 20-fach erhöhten relativen Risiko für Brustkrebs einher. Das bedeutet ein 30 – 60%-iges Risiko mit 60 Jahren an Brustkrebs zu erkranken, verglichen mit 3% in der allgemeinen Bevölkerung [12]. Keimbahnmutationen in BRCA1 oder BRCA2 verleihen ein hohes Lebenszeitrisko für das Mamma- und Ovarialkarzinom. Das Lebenszeitrisko für Brustkrebs wird zwischen 46 – 85% für BRCA1- und 43 – 84% für BRCA2-Mutationsträgerinnen angegeben [14]. Frauen mit der identischen Mutation in einem dieser Gene können jedoch ein unterschiedliches Risikoprofil aufweisen, beeinflusst durch zusätzliche Faktoren. Bei Mutationsträgerinnen wird die Erstdiagnose Brustkrebs durchschnittlich mit 42 Jahren gestellt, etwa 20 Jahre früher als über alle Brustkrebsfälle gemittelt [19]. Mutationen in BRCA1 oder BRCA2 gehen folglich mit einem früh auftretenden Brustkrebs einher. Die Krebsdisposition wird als autosomal dominantes Merkmal in Familien, bei denen eine Mutation segregiert, weitergegeben. BRCA1 und BRCA2 wirken auf zellulärer Ebene jedoch im Sinne der Tumorsuppressoren als rezessive Gene. Mutationen in BRCA1 und BRCA2 begründen etwa 16% des familiären Risikos für Brustkrebs [12].

Obwohl BRCA1 und BRCA2 keine ersichtlichen Sequenzhomologien aufweisen, besitzen sie dennoch eine Reihe von Gemeinsamkeiten [14, 19, 20]. Beide Gene haben eine relativ komplexe genomische Struktur und Größe: BRCA1 besteht aus 22 Exonen und kodiert für ein Protein mit 1.863 Aminosäuren, während BRCA2 27 Exone besitzt und für ein 3.418 Aminosäuren-Protein kodiert. Im Gegensatz zu vielen anderen Tumorsuppressorgenen zeigen sie eine relativ geringe Konservierung zwischen unterschiedlichen Spezies, mit der Ausnahme weniger kleiner Domänen. Beide Gene wirken als Sensoren von DNA-Schäden und sind an Reparaturprozessen beteiligt. Bei ihrer Inaktivierung akkumulieren genetische Defekte und genetische Instabilität entsteht.

Das BRCA1-Protein hat Funktionen in DNA-Reparatur, Ubiquitinierung von Proteinen,

Chromatinumbau und Kontrolle des Zellzyklus [14, 21]. Es kann BRCA2, p53, RAD51 und viele andere Proteine binden, die in den Zellzyklus und die Antwort auf DNA-Schäden involviert sind. BRCA1 wird durch die Proteine ATM und CHEK2 als Antwort auf DNA-Schäden phosphoryliert. Zellen ohne funktionsfähiges BRCA1 gehen nach DNA-Schäden nicht in den G2-Zellzyklusarrest und weisen eine unzureichende Transkriptionsgebundene Reparatur auf [19].

Das BRCA2-Protein ist involviert in die Reparatur von DNA-Doppelstrangbrüchen durch homologe Rekombination und in die Aufrechterhaltung chromosomaler Integrität [21]. Es kann u.a. BRCA1 und RAD51 binden [19]. Biallelische Mutationen in BRCA2 verursachen die seltene Form D1 der Fanconi Anämie (FA-D1) [22]. FA ist eine rezessive Erbkrankheit und ist charakterisiert durch multiple Fehlbildungen, Pigmentanomalien der Haut, progressiven Knochenmarksschwund und eine mit steigendem Alter starke Zunahme bösartiger Erkrankungen, v.a. von Leukämien. Diese Erkrankung ist genetisch heterogen, d.h. durch Mutationen in verschiedenen Genen bedingt, die alle zu chromosomaler Instabilität führen. Es wird jedoch als unwahrscheinlich angesehen, dass Mutationen in anderen FA-Genen eine wichtige Ursache für hochpenetrante Brustkrebsprädisposition sind [14].

Die Prävalenz heterozygoter Träger einer krankheitsverursachenden Mutation ist in der kaukasischen Bevölkerung etwa 1 : 1.000 (0,051%) für BRCA1 bzw. 1 : 750 (0,068%) für BRCA2 [23]. Sie variiert jedoch erheblich zwischen Populationen. In geographisch oder kulturell isolierten ethnischen Gruppen, wie der aschkenasisch jüdischen, isländischen oder finnischen Bevölkerung, sind Gründermutationen in hoher Frequenz zu finden [24].

BRCA1- oder BRCA2-Mutationsträgerinnen haben neben einem erhöhten Risiko für das Brust- und Ovarialkarzinom auch ein erhöhtes Risiko, andere Krebsarten zu entwickeln. Für BRCA1-Mutationsträger wird ein erhöhtes relatives Risiko für das Kolon-, Zervix-, Uterus-, Pankreas- und Prostatakarzinom vermutet. Bei BRCA2-Mutationsträgern und -trägerinnen ist ein erhöhtes relatives Risiko für das männliche Mammakarzinom, das Gallenblasen-, Gallengangs-, Magen-, Pankreas-, Prostata-, Knochen- und Pharynxkarzinom sowie für das maligne Melanom beobachtet worden [14]. Ferner haben Träger einer BRCA1- oder BRCA2-Mutation, die bereits in ihrer Vergangenheit an Brustkrebs erkrankt sind, ein 65%-iges Risiko für ein zweites primäres Mammakarzinom bis zu einem Alter von 70 Jahren [24]. Obwohl BRCA1 und BRCA2 in gleichen zellulären Prozessen wirken, sind Mutationen in den zwei Genen mit einem unterschiedlichen Krebsprofil assoziiert.

### ***STK11/LKB1*** [14, 24]

Das Gen STK11/LKB1 befindet sich auf Chromosom 19p13.3 und kodiert für eine Serin-Threonin-Kinase, die als Tumorsuppressor agiert. Keimbahnmutationen in diesem Gen verursachen das Peutz-Jeghers Syndrom (PJS). Angaben zur Prävalenz reichen von 1 : 8.900 bis 1 : 280.000. Bei PJS-Patienten kann nicht immer eine Mutation in STK11/LKB1 gefunden werden, was für eine heterogene Grundlage der Erkrankung spricht. PJS ist eine autosomal dominante Erkrankung, gekennzeichnet durch gastrointestinale Polyposis und mukokutane pigmentierte Läsionen. Patienten haben ferner ein erhöhtes Risiko für gastrointestinale Tumore, sowie für das Mamma-, Pankreas-, Ovarial-, Uterus-, Zervix-, Lungen- und Hodenkarzinom mit einem Anteil von 37 – 93%, die einen solchen Krebs bis zu einem Alter von 65 Jahren entwickeln. Das Risiko bis zu einem Alter von 65 Jahren an Brustkrebs zu erkranken, liegt zwischen 29 und 54% mit einem durchschnittlichen Alter bei der Diagnose von 45 Jahren.

***PTEN*** [14, 24]

Das Tumorsuppressorgen PTEN ist auf Chromosom 10q23.3 lokalisiert und kodiert für eine dual-spezifische Phosphatase, deren Defekt das Cowden Syndrom (CS) verursacht. Die Prävalenz der Erkrankung wird mit 1 : 300.000 angegeben. CS ist eine seltene autosomal dominante Erkrankung, die durch multiple Hamartome der Haut, der Brust, der Schilddrüse, des Magendarmtrakts und des Zentralnervensystems und einem erhöhten Risiko für das Mamma-, Uterus- und nichtmedulläre Schilddrüsenkarzinom gekennzeichnet ist. Frauen, die eine Keimbahnmutation in PTEN tragen, haben ein Lebenszeitrisko von 25 – 50% für Brustkrebs mit einem durchschnittlichen Alter bei Diagnose der Erkrankung von 38-46 Jahren. In Brustkrebsfamilien, die sonst keine Zeichen des Cowden Syndroms aufweisen, ist jedoch noch keine Mutation im PTEN-Gen gefunden worden.

***TP53/p53*** [14, 24]

Das TP53-Gen befindet sich auf Chromosom 17p13.1 und kodiert für ein Protein, das in die Kontrolle des Zellzyklus, der Apoptose und der DNA-Reparatur involviert ist. Ein Verlust der TP53-Funktion supprimiert einen Mechanismus, welcher gegen die Akkumulation genetischer Alterationen schützt. TP53 fungiert als Tumorsuppressor. Somatische TP53-Mutationen sind sehr häufig in vielen verschiedenen Tumortypen zu finden. Keimbahnmutationen in TP53 sind sehr selten und bedingen das Li-Fraumeni Syndrom (LFS), ein familiäres Krebsdispositions-Syndrom. Die Krebsneigung wird autosomal dominant vererbt. Weltweit sind weniger als 400 Familien mit einer solchen Keimbahnmutation bekannt. Es ist charakterisiert durch multiple primäre Neoplasien im Kindes- und jungen Erwachsenenalter, bevorzugt Weichteilsarkome, Osteosarkome und Brustkrebs, sowie durch ein erhöhtes Risiko für Gehirntumore, Leukämien und das adrenokortikale Karzinom. Eine der häufigsten Krebsarten in einer LFS-Familie ist das Mammakarzinom mit einer geschätzten Penetranz bei Trägern einer TP53-Mutation von 28 – 56% bis zu einem Alter von 45 Jahren. LFS ist besonders assoziiert mit früh auftretendem Brustkrebs mit einem Durchschnittsalter bei Diagnose von 36 Jahren. Da Mutationen in TP53 sehr selten sind, ist dieses Gen nur zu einem geringen Teil für das familiäre Mammakarzinom verantwortlich.

***CDH1/E-cadherin*** [14]

Das Gen CDH1/E-cadherin ist auf Chromosom 16q22.1 lokalisiert. Keimbahnmutationen in diesem Gen sind mit dem vererbten diffusen Magenkrebs Syndrom (HDGC-Syndrom) assoziiert. Es wird autosomal dominant mit unvollständiger Penetranz vererbt. In einer HDGC-Familie haben Frauen, die eine CDH1-Mutation tragen, ein geschätztes kumulatives Risiko für ein diffuses Magenkarzinom von 83% bei einem Alter von 80 Jahren. Das Lebenszeitrisko, ein Mammakarzinom zu entwickeln, wird mit 20 – 40% angegeben.

Keimbahnmutationen in den vier Genen STK11/LKB1, PTEN, TP53 und CDH1 sind sehr selten und werden bei Patientinnen mit Brustkrebs ohne andere klinische Zeichen des jeweiligen Krebsprädispositionssyndroms kaum gefunden.

Mutationen in den beschriebenen sechs Suszeptibilitätsgenen mit hoher Penetranz sind wahrscheinlich maximal für 20% des familiären Risikos der Erkrankung Brustkrebs verantwortlich (vgl. Abbildung 1).

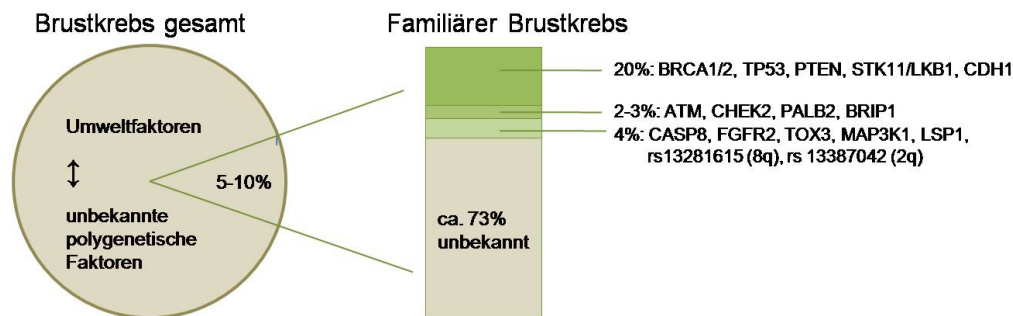


Abbildung 1: **Brustkrebs-suszeptibilität.**

Eine genetische Disposition wird für 5-10% aller Mammakarzinome als Ursache angenommen. Etwa 73% des familiären Brustkrebsrisikos bleiben bis heute ungeklärt.

Bis heute konnte durch genomweite Kopplungsanalysen einer großen Anzahl von Familien, die keine Mutation in BRCA1 oder BRCA2 aufweisen, kein zusätzliches hochpenetranter Brustkrebsgen identifiziert werden. Es wird jedoch vermutet, dass, falls doch weitere solche Gene existieren sollten, diese nur einen kleinen Anteil der familiären Prädisposition begründen. Die verbleibenden etwa 80% des familiären Risikos für Brustkrebs werden den Suszeptibilitätsgenen mit mäßiger Penetranz und den Suszeptibilitätsallelen mit niedriger Penetranz zugerechnet [12].

### Seltene Suszeptibilitätsgene für Brustkrebs mit mäßiger Penetranz

Die Strategie der Identifizierung mäßig penetranter Allele ist die direkte Suche nach krankheitsverursachenden Mutationen in ausgewählten Kandidatengenen. Gute Kandidatengene für Brustkrebs sind beispielsweise solche Gene, die für Proteine kodieren, die in denselben biologischen Signalwegen oder Prozessen interagieren wie BRCA1 und BRCA2. Bis heute konnten mit dieser Methode vier Gene identifiziert werden: ATM, CHEK2, PALB2 und BRIP1 [12].

#### ATM [14]

Das ATM-Gen ist auf Chromosom 11q22-23 lokalisiert. Das ATM-Protein hat eine zentrale Funktion in der Wahrnehmung und Signalisierung von DNA-Doppelstrangbrüchen. Strahlung führt zu einer Autophosphorylierung und damit Aktivierung von ATM. Die ATM-Kinaseaktivität hat viele Substrate, einschließlich TP53, BRCA1 und CHEK2. Homozygote oder compound-heterozygote Keimbahnmutationen im ATM-Gen sind ursächlich für die seltene autosomal rezessive Erkrankung Ataxia teleangiectasia (AT). Die geschätzte Inzidenz für AT liegt zwischen 1 : 40.000 bis 1 : 100.000. AT ist charakterisiert durch zerebelläre Degeneration, Teleangiectasien in Augen und Haut, Immundefizienz, chromosomale Instabilität, erhöhte Sensitivität für ionisierende Strahlung und stark erhöhte Suszeptibilität für Krebs, besonders für Leukämien und Lymphome, aber auch für das Mammakarzinom. Bei heterozygoten Mutationsträgerinnen ist die Einschätzung des Risikos für Brustkrebs kontrovers.

#### CHEK2 [14, 25]

Das CHEK2-Gen befindet sich auf Chromosom 22q12.1. Es kodiert für eine G2 Kontrollpunkt-Kinase, die eine wichtige Rolle in der DNA-Reparatur hat und als Antwort auf DNA-Schädigung durch Phosphorylierung von ATM aktiviert wird. Die Aktivierung von CHEK2 führt zur Phosphorylierung weiterer Schlüsselproteine des Zellzyklus, einschließlich BRCA1 und TP53. CHEK2 wurde als Suszeptibilitätsgen aufgrund der Mutation

CHEK2\* 1100delC identifiziert, welche ein mäßig penetrantes Suszeptibilitätsallel für Brustkrebs darstellt.

### ***PALB2*** [26, 27]

PALB2 kodiert für ein Protein, das mit BRCA2 interagiert und mit dessen nukleärer Lokalisation und Stabilität in Verbindung steht. Das PALB2-Protein wird für einige Funktionen von BRCA2 in der homologen Rekombination und der Reparatur von DNA-Doppelstrangbrüchen benötigt. Biallelische Mutationen in PALB2, ähnlich zu den biallelischen BRCA2-Mutationen, sind für die Form N der Fanconi Anämie verantwortlich (FA-N). Monoallelische krankheitsrelevante Mutationen verursachen ein 2,3-fach erhöhtes Risiko für Brustkrebs.

### ***BRIP1*** [28]

BRIP1 (auch BACH1) kodiert für eine Helikase, die mit BRCA1 interagiert und ebenfalls eine BRCA1-unabhängige Funktion in der DNA-Reparatur und Kontrolle von Checkpoints aufweist. Träger einer monoallelischen inaktivierenden Mutation im BRIP1-Gen tragen eine Prädisposition für Brustkrebs. Das relative Risiko liegt etwa bei 2,0. Biallelische Mutationen in BRIP1 verursachen die Unterform J der Fanconi Anämie (FA-J).

In jedem der vier Gene ATM, CHEK2, PALB2 und BRIP1 sind verschiedene krankheitsverursachende Mutationen gefunden worden, wobei jedoch jede dieser Mutationen sehr selten ist. In Studien unter der allgemeinen UK-Bevölkerung sind 0,6% heterozygot für eine CHEK2-Mutation, etwa 0,4% sind heterozygote Träger einer ATM-Mutation, für BRIP1 oder PALB2 liegt die Fraktion heterozygoter Mutationsträger bei weniger als 0,1% [12]. CHEK2, ATM, BRIP1 und PALB2 haben viele Gemeinsamkeiten mit BRCA1 und BRCA2, gehen jedoch nur mit einem zwei- bis vierfach erhöhten relativen Risiko für Brustkrebs einher. Träger einer Mutation in einem mäßig penetranten Brustkrebsgen haben so ein 6 – 10%-iges Risiko im Alter von 60 Jahren diese maligne Erkrankung zu erleiden. In der Allgemeinbevölkerung liegt das Risiko bei etwa 3%. Nach Schätzungen verursachen Mutationen in diesen vier Genen einen Anteil von 2 bis 3% des familiären Brustkrebsrisikos [12] (vgl. Abbildung 1).

In der Gruppe der seltenen Suszeptibilitätsgene für Brustkrebs mit mäßiger Penetranz kann noch das seltene autosomal dominante Muir-Torre-Syndrom erwähnt werden, das durch Mutationen in den Genen MSH2 oder MLH1 verursacht wird. Es ist vor allem mit multiplen benignen und malignen Talgdrüsentumoren sowie Karzinom innerer Organe, v.a. des Kolons, assoziiert. Patienten weisen ein etwa 2- bis 3-fach erhöhtes Brustkrebsrisiko auf [13].

### **Verbreitete Suszeptibilitätsallele für Brustkrebs mit niedriger Penetranz**

Weit verbreitete Allele, die das Brustkrebsrisiko geringfügig erhöhen, können durch genomweite SNP-Analysen in Assoziationsstudien mit einer großen Anzahl an Fällen und Kontrollen identifiziert werden [12]. Die Frequenz einer genetischen Variante in betroffenen Individuen - den Fällen - wird zu der in Individuen ohne die Erkrankung - den Kontrollen - verglichen. Eine allelische Assoziation ist gegeben, wenn die Genotypenverteilung (Frequenz) unter den Fällen und den Kontrollen verschieden ist. Im Gegensatz hierzu besitzen Kopplungsstudien nicht die statistische Power, Allele mit einem nur moderaten Risiko auf die Erkrankung zu detektieren. Bisher konnten nur wenige statistisch zuverlässige Suszeptibilitätsallele mit niedriger Penetranz gefunden werden, die



auch in anderen Studien repliziert werden konnten. Fünf der gefundenen Risikoallele liegen innerhalb Regionen mit Kopplungsungleichgewicht (linkage disequilibrium, LD) und umfassen proteinkodierende Gene: CASP8 (rs1045485), FGFR2 (rs2981582), TOX3 (rs3803662), MAP3K1 (rs889312) und LSP1 (rs3817198) [29–32]. Eine Assoziation kann durch diese Gene oder auch andere Gene in dieser Region bzw. durch noch nicht bekannte genetische Elemente bedingt sein. Zwei der bekannten Suszeptibilitätsallele, rs13281615 (8q) und rs13387042 (2q), liegen in Regionen mit keinem bekannten proteinkodierenden Gen [29, 31]. Auch wenn eine krankheitsverursachende Variante identifiziert worden ist, ist nicht immer ersichtlich, welches Gen den Effekt vermittelt. Die bisher gefundenen Gene mit niedriger Penetranz unterscheiden sich von den oben beschriebenen dadurch, dass sie keine Rolle in der DNA-Reparatur spielen.

Das relative Brustkrebsrisiko, das mit dem Tragen einer einfachen Kopie eines solchen Risikoallels assoziiert ist, reicht von 1,07 bis 1,26. Die Prävalenz jedes Risikoallels ist mit 28 – 87% sehr hoch. Da sie jedoch nur mit einem niedrigen Brustkrebsrisiko verbunden sind, ist ihr Beitrag zum familiären Brustkrebsrisiko mit etwa 4% ebenfalls gering [12] (vgl. Abbildung 1).

Eine immer noch große Fraktion (>70%) des familiären Brustkrebsrisikos bleibt bis heute ungeklärt. Es wird vermutet, dass ein großer Anteil der familiären Häufung dieser Erkrankung wahrscheinlich durch eine Koinzidenz vieler Suszeptibilitätsallele mit geringem Risiko bedingt ist [12]. In einem polygenetischen Modell sind viele Suszeptibilitätsallele additiv oder multiplikativ, zusammen mit Umweltrisikofaktoren, für Brustkrebs verantwortlich, jedoch nicht mit einem einzelnen Gen, das eher eine große Fraktion der familiären Aggregation ausmacht. Individuen, die wenige solcher Allele tragen, haben im Vergleich zu denjenigen, die viele davon tragen, ein niedrigeres bis allgemeines Populationsrisiko. Ferner kann das mitochondriale Genom ein zusätzliches genetisches Element mit niedriger Penetranz für Brustkrebs darstellen und das individuelle Risiko modifizieren. Die Betrachtung großer Familien mit vielen Fällen, die an Brustkrebs in jungen Jahren erkrankt sind, lässt jedoch auch vermuten, dass weitere hochpenetrante autosomal dominante Gene für Brustkrebs existieren.

### 1.3 Strategien zur Identifizierung von Krankheitsgenen

Es gibt verschiedene Strategien, Krankheitsgene zu identifizieren. Krankheitsgene sind Gene, die Mutationen tragen oder einer veränderten Regulation unterliegen, was zur Entstehung einer Erkrankung führt. Man unterscheidet die funktionsspezifische Klonierung, die positionelle Klonierung, das positionsunabhängige und das positionelle Kandidatengenverfahren [33]. Ferner können Chromosomenanomalien auf die genetische Lokalisation des Defekts hinweisen.

Bei der **funktionsspezifischen Klonierung** wird das Krankheitsgen aufgrund von Informationen über die Funktion des Genprodukts bestimmt. Durch biochemische und zellbiologische Untersuchungen kann ein Protein identifiziert werden, aus dessen Aminosäuresequenz Rückschluss auf die Basensequenz gezogen und das dazugehörige Gen identifizieren werden kann [33].

Das **positionsunabhängige Kandidatengenverfahren** geht von der Frage aus, in welchem bekannten Gen eine Mutation eine bestimmte Krankheit verursachen könnte [3]. Die Erkrankung wird mit phänotypisch ähnlichen Erkrankungen beim Menschen oder Tieren verglichen, für die der genetische Defekt bereits bekannt ist. Kandidatengene stel-

len solche Gene dar, die eine ähnliche Funktion aufweisen oder der gleichen Genfamilie angehören.

Bei der **positionellen Klonierung** und dem **positionellen Kandidatengenverfahren** ist von dem gesuchten Krankheitsgen nur die Zuordnung zu einer chromosomalen Teilregion bekannt [3]. Informationen über die Pathogenese der Erkrankung oder die Funktion des Genprodukts sind nicht notwendig. Die chromosomale Zuordnung kann durch eine Kopplungsanalyse mit anschließender Haplotypenanalyse (vgl. Abschnitt 2.3) oder durch chromosomale Anomalien erhalten werden. Während bei der positionellen Klonierung in einem relativ zeitintensiven Schritt die Kandidatenregion kloniert werden muss und die darin enthaltenen Gene erfasst werden müssen, ist dieser Schritt im positionellen Kandidatengenverfahren wesentlich erleichtert. Durch die Arbeit des “Human Genome Project” sind physikalische und genetische Kartierungen der bisher bekannten menschlichen Gene im Genom verfügbar. Die Gene in der Kandidatenregion können mit Hilfe einer computergestützten Recherche in Genom-Datenbanken (z.B. USCS Genome Browser unter <http://genome.ucsc.edu/>) ermittelt werden. Wichtig ist zu beachten, dass bei diesem Kandidatengenverfahren nur solche Gene analysiert werden können, die bereits bekannt sind.

Unter den Genen, die sich in einer Kandidatenregion befinden, sind solche als Kandidatengene in Erwägung zu ziehen, die aufgrund ihrer Funktion oder einem entsprechenden Expressionsmuster bei Mutation den Phänotyp der Erkrankung hervorrufen könnten. Eine Beweismöglichkeit, dass ein Kandidatengene tatsächlich das gesuchte Gen ist, ist die Suche nach patientenspezifischen Mutationen in der Gensequenz. Die Auswirkung einer möglichen pathogenen Mutation auf die Erkrankung muss jedoch durch weitere funktionelle Studien bestätigt werden.

Das positionelle Kandidatengenverfahren ist in dieser Arbeit die gewählte Strategie, um bei der untersuchten Brustkrebs-Familie das verantwortliche Gen zu finden.

Die folgende Abbildung zeigt die angewendete Vorgehensweise:

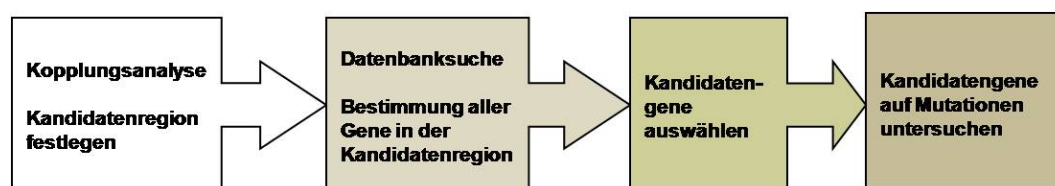


Abbildung 2: *Logik des positionellen Kandidatengenverfahrens.*

## 1.4 Fragestellung der vorliegenden Arbeit

In der vorliegenden Arbeit sollte mittels des positionellen Kandidatengenverfahrens und einer größeren Familie, in der das Mammakarzinom in einer autosomal dominanten Weise segregiert, nach einem neuen “Brustkrebs-Gen” gesucht werden. Im Vorfeld wurden bereits Mutationen in den Genen BRCA1 und BRCA2, die vor allem ursächlich für das familiäre Mamma- und Ovarial-Syndrom sind, ausgeschlossen. Bei einer daraufhin durchgeführten genomweiten Kopplungsanalyse mit dem GeneChip Human Mapping 50K XbaI (Affymetrix) war unter anderem ein Bereich auf Chromosom 17q22 auffällig, in dem das BRIP1-Gen lokalisiert ist. Die anschließende Sequenzierung dieses bekannten Suszeptibilitätsgens zeigte jedoch keine pathogene Veränderung.

Eine detaillierte Auswertung der Kopplungsdaten unter Annahme von sporadischen Phänotypen bei ein oder zwei betroffenen Familienangehörigen weisen auf acht Kandidatenregionen in acht unterschiedlichen chromosomalen Regionen hin. In jeder dieser Regionen könnte das gesuchte Krankheitsgen mit einer vernünftigen Wahrscheinlichkeit zu finden sein.

Im Folgenden wurde die Kopplung der acht Kandidatenregionen mit der Erkrankung Brustkrebs mittels einer Genotypisierung von Mikrosatelliten-Markern bestätigt und die genauen Grenzen der Kopplungsbereiche bestimmt. Nach Identifizierung der darin liegenden Kandidatengene wurde mittels direkter Sequenzanalyse nach brustkrebsassoziierten Mutationen in einigen ausgewählten Genen gesucht.

## 2 MATERIAL, PATIENTEN UND METHODEN

### 2.1 Verwendete Geräte, Software und Reagenzien

#### 2.1.1 Geräte und Software

Gerät	Modell	Hersteller/Quelle
DNA-Thermocycler	T1-Thermoblock	Biometra®
DNA-Thermocycler	T-Gradient Thermoblock	Biometra®
DNA-Thermocycler	T3000-Thermoblock	Biometra®
Dunkelhaube	biostep Dark Hood DH-30/32	biostep GmbH
Dunkelhaube Software	Argus X1 Version 4.0.81	biostep GmbH
Dunkelhaube Drucker	Digital Monochrome Printer P93DW	Mitsubishi Electric
Sequenzier	3130xl Genetic Analyser	Applied Biosystem
Genotyper Software	GeneMapper® Software Version 3.7	Applied Biosystem
Sequenzier Software	Sequencing Analysing Software Version 5.2 Patch2	Applied Biosystem
	SeqMan <sup>TM</sup> II expert sequence analysis software	DNASTAR
	Chromas Version 2.23	Technelysium Pty Ltd
Spektralphotometer	NanoDrop® ND-1000 Spectrophotometer	NanoDrop Technologies
Spektralphotometer Software	NanoDrop ND-1000 Version 3.5.2	Coleman Technologies Inc.
Vortexer	VORTEX-GENIE2®	Scientific Industries Inc.
Tischzentrifuge	Mini Centrifuge	Qualitron
	Mini Centrifuge	Labnet International Inc.
Zentrifuge	Megafuge 1.0R	Heraeus
Laborwage	SBC 52	Scaltec Instruments
Pipetten	Eppendorf Reference 10, 100, 1000	Eppendorf
	Eppendorf Multipipette® plus	Eppendorf
Elektrophoresekammer mit Zubehör	Blue Marine 200	Serva Electrophoresis GmbH
Netzgerät	Blue Power Plus BP-Plus	Serva Electrophoresis GmbH
Haplotypen Software	HaploPainter Version 029-5	<a href="http://haplopainter.sourceforge.net">http://haplopainter.sourceforge.net</a>
Editor	ApE-A plasmid Editor Version 1.12	<a href="http://www.biology.utah.edu/jorgensen/wayned/ape">http://www.biology.utah.edu/jorgensen/wayned/ape</a>

## 2.1.2 Reagenzien

Reagenz	Beschreibung	Hersteller
dNTP Set: dATP, dCTP, dGTP, dTTP	100 mM dNTP-Lösung	metabion international AG
H <sub>2</sub> O ultrasteril	Water Purification Systems	Millipore
10× PCR Puffer	0,5 M KCl 0,2 M Tris-HCl (pH 8,3) 15 mM MgCl <sub>2</sub> ad 30 ml H <sub>2</sub> O	/
h-Taq	am eigenen Institut exprimierte und gereinigte Taq-Polymerase	/
DMSO	Dimethylsulfoxide 1ml	STRATAGENE®
5× Q-Solution	Q-Solution, 5x konzentriert	QIAGEN
5× AccoPrime™ GC-Rich DNA Polymerase	2 U/µl AccoPrime™ GC-Rich DNA Polymerase	invitrogen™
5× Buffer A	300 mM Tris-HCl (pH 9.2) 10 mM MgSO <sub>4</sub> 150 mM NaCl 1 mM dGTP 1 mM dATP 1 mM dTTP 1 mM dCTP thermostabile AccuPrime™ Proteine Enhancers	invitrogen™
GeneScan™ 500 Rox™ Size Standard	8 nM, enthält 16 einzelsträngige Fragmente (35, 50, 75, 100, 139, 150, 160, 200, 250, 300, 430, 350, 400, 450, 490 und 500 Basen)	Applied Biosystem
HiDi™ Formamide	hoch deionisiertes Formamid	Applied Biosystem
Exonuclease I	10 U/µl Exonuclease I	USB Corporation
Antarctic Phosphatase	5,000 U/ml Antarctic Phosphatase	New England BioLabs® Inc.
BigDye® Terminator v1.1/ Sequencing Standard Kit	Kit beinhaltet u.a. fertigen Reaktions-Mix und 5× BigDye Sequencing Buffer v1.1/v.3.1	Applied Biosystem
Ethanol 100%	Ethanol p.a. 100%	Merck
Agarose	Biozym LE Agarose	Biozym Scientific und Biozym Biotech Trading GmbH
Ethidiumbromid	Ethidiumbromid-Lösung 0,07%	AppliChem GmbH

5× TBE	54,0 g TRIS 27,5 g Boric acid 20 ml 0,5M EDTA-Solution (pH 8.0) ad 1000 ml H <sub>2</sub> O	/
1 Kb Plus DNA-Leiter	1 µg/µl, enthält 20 doppelsträngige DNA- Fragmente von 100 bp bis 12 kb	invitrogen <sup>TM</sup>

## 2.2 Patienten

Der Stammbaum der Familie, die in dieser Studie molekulargenetisch untersucht wurde, ist in Abbildung 3 dargestellt. Die Familie besteht aus vier Generationen (I - IV) und zeigt mit sieben Frauen (1-3, II-10, III-1, III-5, III-6, III-7 und III-9), die an Brustkrebs erkrankt sind, eine deutliche Häufung dieser Krebsart in der Familie. Die Erstdiagnose Brustkrebs wurde bei mindestens sechs der sieben Frauen (II-10, III-1, III-5, III-6, III-7 und III-9) vor dem fünfzigsten Lebensjahr gestellt. Es liegt somit eindeutig ein familiäres Mammakarzinom vor. In die molekulargenetische Untersuchung konnten elf Personen (II-5, II-6, II-7, II-10, II-12, II-13, III-1, III-5, III-6, III-7 und III-9) aus den Generationen III und IV einbezogen werden. Weitere Familienmitglieder waren bereits verstorben oder standen für die Untersuchung nicht zur Verfügung. Person II-10 ist ebenfalls bereits verstorben, es stand jedoch eine DNA-Probe für die molekulargenetische Untersuchung zur Verfügung. Ferner wurden die Kinder in fünfter Generation (nicht im Stammbaum dargestellt) nicht berücksichtigt, da in dieser Generation aufgrund des jungen Alters keine prädiktiven Aussagen gemacht werden sollten.

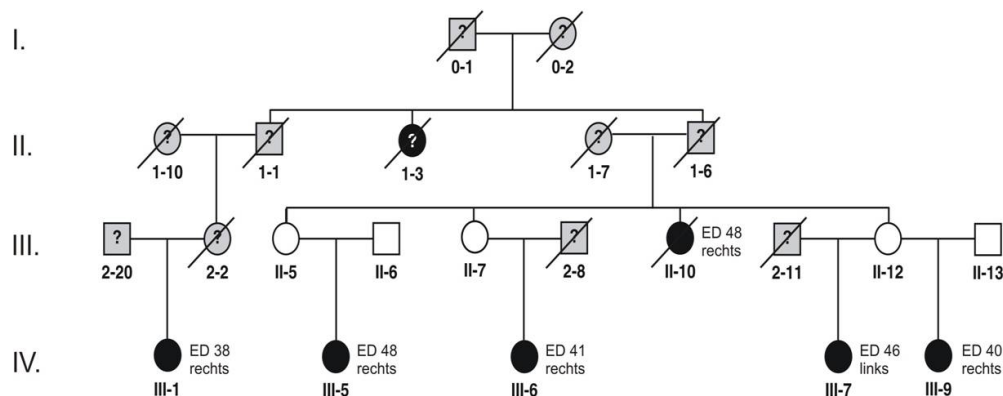


Abbildung 3: **Ausschnitt aus dem Stammbaum der untersuchten Familie.**

Zeichenerklärung: männlich ( $\square$ ), weiblich ( $\circ$ ); an Brustkrebs erkrankt ( $\bullet$ ), nicht an Brustkrebs erkrankt ( $\square, \circ$ ); grau hinterlegte Personen mit Fragezeichen standen für die Analyse nicht zur Verfügung; Angabe des Alters bei Erstdiagnose (ED) Brustkrebs, Seite der betroffenen Brust (links, rechts). Mit einem Querstrich versehene Symbole kennzeichnen verstorbene Familienmitglieder.

Die Analyse des Stammbaums lässt auf eine autosomal dominante Vererbung des Gendefektes in der Familie schließen, der die Erkrankung Brustkrebs bedingt. Da jedoch in

dritter Generation die Frauen II-5, II-7 und II-12 nicht bzw. noch nicht den Phänotyp Mammakarzinom aufweisen, wirkt sich der Defekt mit unvollständiger Penetranz aus. Diese drei Frauen sind obligate Carrier, d.h. Überträgerinnen des Gendefekts. Ein X-chromosomal gebundener Vererbungsmodus erweist sich in der Stammbaumanalyse als unwahrscheinlich. Der X-gekoppelte Gendefekt hätte über die Männer 1-1 und 1-6 weitergegeben werden müssen. Bei diesen ist jedoch keine Erkrankung Brustkrebs bekannt.

Im Vorfeld der Arbeit wurde bereits aus peripheren Blutproben der elf zu untersuchenden Familienmitglieder DNA nach einem Standard-Protokoll isoliert. Für die Verwendung in der Kopplungsanalyse mit Mikrosatelliten-Markern und in der Sequenzierung wurde die DNA mit ultrasterilem H<sub>2</sub>O auf eine Konzentration von 40ng/µl verdünnt.

## 2.3 Die Methode der familiären Kopplungsanalyse

### 2.3.1 Die Kopplungsanalyse

#### Ziel der Kopplungsanalyse [3]

Die Kopplungsanalyse dient der Untersuchung der Kosegregation einer bestimmten Erkrankung mit einer bestimmten chromosomalen Region. Sie ist der erste Schritt im Rahmen des positionellen Kandidatengenverfahrens (vgl. Abschnitt 1.3). Die Analyse bedarf großer Familien mit einer familiär gehäuft auftretenden Erkrankung. Die in die Untersuchung einbezogenen Individuen müssen phänotypisch genau charakterisiert sein. Durch eine Stammbaumanalyse können zusätzliche obligate Genträger identifiziert werden.

Für die Kopplungsanalyse werden DNA-Marker genotypisiert, die über das gesamte Genom verteilt sind. Ein Marker ist ein beliebiges Merkmal, mit dessen Hilfe man einen Chromosomenabschnitt in einem Stammbaum verfolgen kann. Er liegt in jedem Individuum in zwei oder mehreren Allelen vor, die nach den Mendelschen Regeln vererbt werden. Die verwendeten genetischen Marker sollen zwei Kriterien erfüllen: zum einen sollen sie ausreichend polymorph sein, zum anderen einen hohen Heterozygotiegrad aufweisen, d.h., eine zufällig ausgewählte Person sollte mit hoher Wahrscheinlichkeit heterozygot und damit informativ für diesen Marker sein. Für eine Kopplungsanalyse werden heutzutage entweder hochinformativ Mikrosatelliten-Marker (vgl. Abschnitt 2.3.3) oder Einzelnukleotidpolymorphismen (vgl. Abschnitt 2.3.2) eingesetzt. Es sollen diejenigen DNA-Marker herausgefiltert werden, die mit dem Gen gekoppelt sind, dessen Defekt die familiär bedingte Erkrankung verursacht. Je nach zugrundeliegendem Vererbungsmodell wird beispielsweise bei Vorliegen einer autosomal dominanten Vererbung nach Allelen gesucht, die alle Genträger in einfacher Anzahl besitzen, die Gesunden jedoch nicht. Solche Genträger-spezifischen Allele können mit dem genetischen Defekt gekoppelt sein, d.h. sie segregieren mit der Erkrankung. Da seit der Arbeit des "Human Genome Projects" die genaue Lokalisation der DNA-Marker bekannt ist, können die chromosomalen Regionen bestimmt werden, bei denen eine Kopplung mit der Erkrankung vorliegt.

#### Mathematische Grundlage der Kopplungsanalyse [3]

Wenn zwei Genorte auf verschiedenen, nicht homologen Chromosomen liegen, beobachtet man freie Rekombination zwischen ihnen. Unter den Nachkommen erwartet man einen gleichen Anteil an rekombinierten und nicht-rekombinierten Allelen. Befinden sie sich jedoch auf dem gleichen Chromosom, so werden sie häufiger gemeinsam vererbt. Dieses Phänomen wird als Genkopplung bezeichnet. Je näher zwei DNA-Sequenzen auf einem Chromosom nebeneinander liegen, desto häufiger werden sie gekoppelt vererbt. Umge-

kehrt nimmt die Wahrscheinlichkeit einer Rekombination durch ein Crossing-over ab, je näher zwei Loci auf einem Chromosom beisammen liegen. Das Ausmaß genetischer Kopplung wird durch die Rekombinationsfraktion  $\theta$  (theta) beschrieben. Die Rekombinationsfraktion gibt die Wahrscheinlichkeit an, dass ein von einem Elternteil erzeugter Gamet rekombinant ist [34]. Folglich sind Gene, die unabhängig voneinander segregieren, nicht gekoppelt und weisen eine Rekombinationsfraktion von  $\theta = \frac{1}{2}$  auf. Gekoppelte Gene zeichnen ein  $\theta < \frac{1}{2}$  aus. Bei vollständiger Kopplung ist  $\theta$  gleich null.

Die Rekombinationsfraktion wird als Maß für die Entfernung zwischen zwei Loci verwendet und dient der Erstellung genetischer Karten. Die Abstände von Genen werden in centi-Morgan (cM) angegeben. Zwei Loci, die 1% Rekombination (d.h.  $\theta = 0,01$ ) zeigen, haben einen genetischen Abstand von 1cM. Da die Rekombinationsfraktion maximal 0,5 erreichen kann, kann  $\theta$  nicht einfach auf genetischen Karten addiert werden. Kartierungsfunktionen beschreiben die mathematische Beziehung zwischen der Rekombinationsfraktion  $\theta$  und dem Abstand  $\omega$  auf einer genetischen Karte.

Die Funktion nach Haldane geht davon aus, dass Crossing-over über das gesamte Genom zufällig entstehen und sich gegenseitig nicht beeinflussen:  $\omega = -\frac{1}{2} \ln(1 - 2\theta)$ . Eine Kartierungsfunktion, die berücksichtigt, dass die Bildung eines Crossing-over nicht zufällig erfolgt und die Entstehung eines zweiten in seiner Nähe hemmt (Interferenz), ist die Funktion nach Kosambi:  $\omega = \frac{1}{4} \ln[(1 + 2\theta)/(1 - 2\theta)]$ . In einer physikalischen Karte entspricht 1cM durchschnittlich 1Mb (Megabase). Die Beziehung variiert jedoch erheblich in Abhängigkeit vom chromosomalen Locus und Geschlecht.

### Auswertung der Kopplungsanalyse [3]

Die Bewertung von Kopplungsanalysen erfolgt statistisch und computergestützt. Grundlage ist die Abschätzung von  $\theta$  und die Untersuchung der Hypothese von freier Rekombination ( $\theta = \frac{1}{2}$ ) versus Kopplung ( $\theta < \frac{1}{2}$ ).

#### Die Zweipunkt-Analyse

Bei der Zweipunkt-Analyse wird jeder eingesetzte DNA-Marker (Locus 1) auf Kopplung mit einem in der Nähe liegenden hypothetischen Krankheitslocus (Locus 2) untersucht. Man geht von den Möglichkeiten aus, dass die beiden Loci entweder freie Rekombination (Nullhypothese  $H_0$ ) oder Kopplung (Hypothese  $H_1$ ) aufweisen. Die Hypothesen sind charakterisiert durch den Wert der Rekombinationsfraktion; für freie Rekombination ist  $\theta$  folglich  $\frac{1}{2}$  und für Kopplung ist  $\theta < \frac{1}{2}$ .

- $H_0$ :  $\theta = \frac{1}{2}$
- $H_1$ :  $\theta < \frac{1}{2}$

Um die beiden Hypothesen miteinander vergleichen zu können betrachtet man den Quotienten der Wahrscheinlichkeiten (*likelihoods*,  $L$ ) beider Annahmen. Diesen Quotienten bezeichnet man als *likelihood ratio* bzw. *odds ratio*  $R$  [34].

$$R = \frac{\text{Wahrscheinlichkeit, dass Locus 1 und Locus 2 gekoppelt sind mit } \theta < \frac{1}{2}}{\text{Wahrscheinlichkeit, dass Locus 1 und Locus 2 nicht gekoppelt sind mit } \theta = \frac{1}{2}} = \frac{L(H_1)}{L(H_0)}$$

Dieses Wahrscheinlichkeitsverhältnis sagt etwas über die Chance für das Vorliegen einer Kopplung aus. Je größer der Wert für  $R$  ist, desto wahrscheinlicher ist die Annahme der Hypothese  $H_1$ , die eine gekoppelte Vererbung beschreibt, im Vergleich zur freien Rekombination ( $H_0$ ). Als Maß für das Vorhandensein einer Kopplung versus Abwesenheit einer



Kopplung werden üblicherweise sogenannte LOD-Werte ( $\text{LOD} = \text{logarithm of the odds}$ ) verwendet, die als  $Z(\theta)$  angegeben werden. Der LOD ist der Logarithmus zur Basis zehn der *likelihood ratio*  $R$  [34]:

$$Z(\theta) = \log_{10} R = \log_{10} \left[ \frac{L(\theta)}{L(\frac{1}{2})} \right] \quad (0 \leq \theta \leq \frac{1}{2})$$

Mit dieser Gleichung kann für jeden in der Kopplungsanalyse eingesetzten Marker die Rekombinationsfraktion  $\theta$  ermittelt werden, für die  $Z(\theta)$  maximal wird. Die Berechnung des LOD erfolgt für eine Reihe von  $\theta$ -Werten zwischen 0 und  $\frac{1}{2}$  mit Hilfe spezieller Computerprogramme.

Der Wert von  $Z_{\max}$  lässt eine Aussage über eine mögliche Kopplung des Markers mit dem Krankheitsloкус zu. Ein positiver LOD deutet auf Kopplung hin, ein negativer Wert dagegen auf Abwesenheit von Kopplung. Bei einem LOD-Wert von  $X$  an einer bestimmten Region des Genoms ist die Wahrscheinlichkeit, die gegebenen Daten zu erhalten, wenn  $H_1$  zutrifft,  $10^x$  mal größer, als diese Daten zu erhalten, wenn  $H_0$  zutrifft. So ist beispielsweise bei einem  $Z(\theta) = 3$  diese Wahrscheinlichkeit  $10^3 = 1.000$  –mal größer. Wenn  $Z_{\max}$  einen bestimmten kritischen Wert  $Z_0$  erreicht, soll dieser ein signifikanter Beweis für eine Kopplung sein. Nach Morton (1955) [35] spricht ein  $Z_{\max} = +3$  auf autosomalen Loki bzw. ein  $Z_{\max} = +2$  auf X-chromosomalen Loki signifikant für eine Kopplung. Bei einem  $Z = -2$  kann eine Kopplung ausgeschlossen werden. Werte zwischen -2 und +3 lassen keine signifikante Aussage zu. Sie können auf eine Kopplung hinweisen, sollten jedoch durch weitere Marker oder durch Einbeziehung zusätzliche Familienmitglieder in die Analyse bestätigt werden.

Jede voll informative Meiose in einer Familie trägt  $\log(2) \approx 0,3$  zur Höhe des LOD-Wertes bei [34]. Somit ist für ein signifikantes Ergebnis einer Kopplungsanalyse eine ausreichend große Anzahl betroffener Individuen in einer Familie nötig.

### **Die Mehrpunktanalyse** [36]

Eine Mehrpunktanalyse ermöglicht im Vergleich zur Zweipunkt-Analyse die Signifikanz der Kopplungsanalyse zu steigern. Eine Mehrpunktanalyse benötigt eine genetische Karte, die Rahmenkarte, in der die verwendeten Marker einer genauen Position zugeordnet sind. In der vorliegenden Arbeit wurde für die Analyse das Programm Merlin verwendet (vgl. Abschnitt 2.3.2). Ein Vorteil der Multimarkerkartierung ist, dass Probleme vermieden werden können, die durch einen geringen Informationsgehalt von Markern bedingt sind. Der Krankheitsloкус wird bestimmt, indem für jede mögliche Position im Genom anhand der Stammbaumdaten die Gesamtwahrscheinlichkeit für eine Kopplung berechnet wird. Das Ergebnis lässt sich als Kurve darstellen, bei der die Wahrscheinlichkeiten gegen die Position in der genetischen Karte aufgetragen sind. Auf der Y-Achse ist gewöhnlich der LOD-Wert aufgetragen, auf der X-Achse die Multimarkerkarte. Der höchste Punkt des Graphen markiert die wahrscheinlichste Position des Krankheitsloкус. Wenn die Kurve über einen bestimmten chromosomalen Abschnitt hinweg unterhalb  $Z \leq -2$  bleibt, kann eine Kopplung des Krankheitsloкус mit diesem Chromosomenabschnitt ausgeschlossen werden.

Bei der Mehrpunktanalyse wird gelegentlich der sogenannte Positionswert (*location score*) benutzt. Das ist der zweifache natürliche Logarithmus der *likelihood ratio*  $R$ . Den Zusammenhang zwischen dem LOD-Wert und dem Location-Wert beschreibt folgende Gleichung:

$$\text{location score} = \text{LOD} \cdot 2 \ln(10) \approx \text{LOD} \cdot 4,6 \text{ [36].}$$

### 2.3.2 Kopplungsanalyse mit 50K XbaI-GenChip

#### Einzelnukleotidpolymorphismen

Einzelnukleotidpolymorphismen (*single nucleotide polymorphisms*, SNP) bestehen in der Regel aus zwei Allelen, sind also weniger polymorph und informativ als Mikrosatelliten-Marker. Sie ermöglichen jedoch eine Genotypisierung mit sehr hoher Markerdichte und die Analyse durch automatisierte Verfahren in einem großen Maßstab. Einzelnukleotidpolymorphismen haben eine hohe Dichte im Genom von durchschnittlich einem SNP pro Kilobase DNA [3]. Es sind zurzeit über 6 Millionen geprüfte humane SNPs in öffentlichen Datenbanken hinterlegt. Sie ermöglichen die bisher größte Feinauflösung in der Kartierung bei Kopplungsanalysen.

#### Genotypisierung mit 50K XbaI-GenChip

Die genomweite Genotypisierung der elf zur Verfügung stehenden Familienmitglieder (II-5, II-6, II-7, II-10, II-12, II-13, III-1, III-5, III-6, III-7 und III-9) erfolgte mit dem GeneChip® Human Mapping 50K Array XbaI 240 von Affymetrix. Die Analyse wurde am Institut für Funktionale Genetik und Genomik des Max-Delbrück-Center für Molekulare Medizin in Berlin (AG Hübner) unter Mitarbeit von Dr. Franz Rüschenhoff durchgeführt. Der verwendete 50K Array XbaI 240 diente der Genotypisierung von 58.960 SNPs. Der Median des physikalischen Abstands zwischen den SNPs beträgt 8,5 kb, der durchschnittliche Abstand 23,6 kb. Die durchschnittliche Heterozygotie liegt bei 0,30 [37]. Die Genotypen wurden mit dem BRLMM-Algorithmus gecallt, der eine hohe Callrate erzeugt und eine hohe Genauigkeit bzw. Fehlerfreiheit aufweist [38].

#### Auswertung mit Merlin

Die statistische Auswertung der Genotypisierungsdaten erfolgte mit dem Unix-basierten Programm Merlin bzw. MINX (Merlin für X) und wurde von Dr. Franz Rüschenhoff durchgeführt. Merlin arbeitet mit dem Lander-Green-Algorithmus, der ermöglicht, eine beliebige Anzahl von Multimarkerdaten in Stammbäumen mittlerer Größe zu analysieren [39].

Als Input-Dateien wurden benötigt:

- Stammbaum-Datei (*pedigree file*): Beschreibung des Geschlechts, der Beziehung zwischen den Individuen, des Phänotyps (nicht betroffen, betroffen, fehlender Phänotyp) und der Genotypen der analysierten Marker. Ein fehlender Phänotyp wurde für alle Personen, die für die Untersuchung nicht zur Verfügung standen und nicht an Brustkrebs erkrankt waren, angegeben.
- Einzelheiten-Datei (*data file*): Beschreibung der Stammbaumdatei. Dem Phänotyp wird eine Krankheit, den Genotypen ein Markername zugeordnet.
- Genetische Rahmenmarker-Datei (*map file*): Beschreibung von Chromosom, Markername und Position (in cM) für jeden analysierten Marker.
- Allelfrequenz-Datei: Beschreibung der Allelfrequenz für jedes Allel der untersuchten Marker.

Die Analysen wurden für die in dieser Arbeit untersuchte Familie unter zwei Annahmen durchgeführt:

**Annahme 1**

Allen Personen, die an Brustkrebs erkrankt sind, liegt ein genetischer Defekt zugrunde, der in der Familie bei allen Betroffenen mit der Erkrankung segregiert. Alle erkrankten Individuen wurden in der Stammbaum-Datei als betroffen angegeben.

**Annahme 2**

Mindestens eine betroffene Person erkrankte an Brustkrebs nicht aufgrund des in der Familie mit der Erkrankung segregierenden genetischen Defekts, sondern aufgrund sporadischer Genese. Sie ist keine Trägerin des krankheitsverursachenden Allels, sondern lediglich eine Phänokopie. Annahme 2 ist in Hinblick auf die hohe Prävalenz von Brustkrebs in der Bevölkerung zu beachten (vgl. Abschnitt 1.2.1). In der Stammbaum-Datei wird diese Person als fehlender Phänotyp (*unknown*) angegeben. Bei allen anderen betroffenen Individuen wird die Segregation der Erkrankung mit einem familiär bedingten genetischen Defekt angenommen. Die Berechnung erfolgte, indem nacheinander die Betroffenen III-1, III-5, III-6, III-7, III-9 und II-10 auf “unknown” gesetzt wurden.

Mit Merlin können unter anderem parametrische und nicht-parametrische Kopplungsanalysen, Fehlerdetektion und Haplotypenanalysen durchgeführt werden. Im Vorfeld der Analyse sind diejenigen SNPs ausgeschlossen worden, deren Genotypen einen Mendelfehler in der Familie zeigten oder als unwahrscheinlich detektiert wurden. Insgesamt wurden 35.642 SNPs über alle Chromosomen verteilt in die Auswertung miteinbezogen.

**Die parametrische Kopplungsanalyse**

Bei der Standard-LOD-Wert-Analyse bzw. parametrischen Analyse ist die Kenntnis über das genaue genetische Modell erforderlich. Dies beinhaltet den Vererbungsmodus, die Genhäufigkeit und die Penetranz jedes Genotyps. Bei der Penetranz ist zu beachten, dass auch nichterkrankte Personen möglicherweise nichtpenetrante Genträger oder erkrankte Personen Phänokopien sein können [3]. Für den Nachweis einer Kopplung darf die Penetranz jedoch auch nicht zu niedrig festgelegt werden, da dies eine zu ungenaue Hypothese testet. Bei dem Vorliegen einer mendelnden Krankheit ist die parametrische Analyse die bevorzugte Berechnung.

Es wurde für jede beschriebene Annahme eine parametrische Kopplungsanalyse mit verschiedenen genetischen Modellen durchgeführt. Die besten Resultate konnten mit einem dominanten Vererbungsmodus (*Rare Dominant*) erzielt werden. Die Frequenz des Krankheitsallels in der Bevölkerung ist auf 0,001, die Wahrscheinlichkeit betroffen zu sein für Individuen, die keine, eine bzw. zwei Kopien des Krankheitsallels tragen (Penetranz) auf 0,00, 0,90 bzw. 0,90 festgelegt worden.

In der folgenden Tabelle sind die Parameter des dominanten Vererbungsmodells aufgeführt:

Frequenz des Krankheitsallels	0,001
Penetranz für Träger von keinem Krankheitsallel	0,00
Penetranz für Träger von 1 Krankheitsallel	0,90
Penetranz für Träger von 2 Krankheitsallelen	0,90

Alle in dieser Arbeit dargelegten Ergebnisse wurden mit diesem Modell berechnet.

In der Output-Datei der parametrischen Kopplungsanalyse wurde jeder Markerposition in der Multimarkerrahmenkarte ein LOD-Wert, der sogenannte *multipoint* LOD-Wert (mLOD), zugeordnet. Als mögliche Kandidatenregionen für eine Kopplung wurden bei der Analyse unter Annahme 2 solche chromosomalen Bereiche in Erwägung gezogen, die einen mLOD-Wert  $> 1,4$  auf den Autosomen erzielten. Das X-Chromosom wurde bei der weiteren Analyse unter dieser Annahme aufgrund des vermuteten autosomal dominant vererbten Gendefekts nicht explizit untersucht (vgl. Abschnitt 2.2). Für die Bewertung der Ergebnisse unter Annahme 1 wurden keine derartigen Einschränkungen getroffen. Hier wurde auch das X-Chromosom wegen eines relativ hohen mLOD-Werts von 1,8 in die weitere Analyse einbezogen.

### **Die nicht-parametrische Kopplungsanalyse**

Die modellfreie bzw. nicht-parametrische Kopplungsanalyse ist die Methode der Wahl zum Suchen nach Allelen oder Segmenten, die betroffene Personen gemeinsam haben. Die Kenntnis über das genaue genetische Modell ist nicht erforderlich. Diese Analyse kann auf eine Familie oder auch ganze Populationen angewandt werden und dient der Kopplungsanalyse komplexer bzw. nichtmendelnder Merkmale. Es erfolgt eine Unterscheidung zwischen Segmenten, die abstammungsidentisch (*identical by descent*, IBD) oder eigenschaftsidentisch (*identical by state*, IBS) sind [3]. Die Analyse kann mit IBD- oder IBS-Daten durchgeführt werden. Die IBD-Analyse ist effektiver, erfordert jedoch die Untersuchung von verwandten Personen.

In der untersuchten Familie wurde zusätzlich für jede Annahme eine nicht-parametrische Kopplungsanalyse durchgeführt. Es wurde nach Chromosomensegmenten gesucht, in denen die erkrankten Personen mit einer höheren Wahrscheinlichkeit übereinstimmen, als zufällig zu erwarten wäre. Die gemeinsamen IBD-Allele wurden bei allen erkrankten Individuen des Stammbaums mit der Nullhypothese einer einfachen mendelnden Vererbung verglichen. Das Ergebnis sind nicht-parametrische LOD-, NPL- bzw. NPLD-Werte. In der Output-Datei von Merlin werden diese als  $Z_{mean}$  bezeichnet.

### **Haplotypenanalyse**

Eine Gruppe gemeinsam vererbter Allele benachbarter Marker auf einem Chromosom wird als Haplotyp bezeichnet. Haplotypen, die zusammen mit einer Erkrankung segregieren und sich nur bei Betroffenen und Genträgern der Familie finden, bezeichnet man als krankheitsspezifische Haplotypen bzw. Risikohaplotypen [3].

Merlin kann unter Verwendung der Informationen des Genflusses in dem vorliegenden Stammbaum die wahrscheinlichsten Haplotypen für die Individuen der Familie rekonstruieren. Es handelt sich jedoch um geschätzte Haplotypen, die in der Realität durchaus anders aussehen können. In der Output-Datei erhält man für jedes Individuum eine Liste der zwei Haplotypen, entsprechend dem mütterlichen und väterlichen Haplotyp. Die Lokalisation wahrscheinlicher Rekombinationsereignisse wird durch zusätzliche Zeichen angezeigt.

### **Bestimmung der flankierenden Marker der Kopplungsregionen**

Durch die Haplotypenanalyse kann der krankheitsspezifische Risikohaplotyp auf den kleinsten gemeinsamen Haplotypen aller betroffenen Personen eingegrenzt werden. Der Nachweis von Rekombinationen zielt auf eine maximale Eingrenzung des Risikohaplotypen ab. Rekombinationsereignisse ermöglichen die Bestimmung von denjenigen Markern, die den möglichen Krankheitsloкус zentromerisch und telomerisch eingrenzen. Das sind

die flankierenden Marker der Kopplungsregionen.

Die initiale graphische Darstellung der Haplotypen erfolgte mit dem Programm HaploPainter Version 029.5<sup>1</sup>. Anschließend wurden die Haplotypen aller elf in die Untersuchung einbezogenen Personen in Excel-Tabellen nebeneinander verglichen. Dies ermöglichte, die informativen Marker herauszufiltern und flankierende Marker für chromosomale Regionen, die Kandidatenregionen für eine Kopplung darstellen, zu bestimmen. Auf diese Weise wurden für jede Kandidatenregion die informativen flankierenden Marker bestimmt, die diese maximal eingrenzen. Zusätzlich erfolgte für jeden flankierenden Marker die Bestimmung des folgenden informativen Markers innerhalb der Kopplungsregion. Diese Marker werden im weiteren Verlauf der Arbeit als "folgende Marker" bezeichnet. Die Region zwischen dem flankierenden und seinem folgenden Marker gibt die Lokalisation der stattgefundenen Rekombination an. Diese kann durch eine Analyse zusätzlicher Marker noch weiter eingegrenzt werden.

### 2.3.3 Bestätigung und weitere Eingrenzung der Kopplungsregionen mit Mikrosatelliten-Markern

In der zusätzlich durchgeführten Genotypisierung von Mikrosatelliten-Markern wurden die in der Haplotypenanalyse herausgearbeiteten krankheitsspezifischen Haplotypen genauer untersucht. Ziel war es, falsch positive Kopplungen aufzudecken und die tatsächlich gekoppelten Regionen eventuell weiter einzugrenzen.

#### Mikrosatelliten-Marker [3, 33]

Eine Mikrosatelliten-DNA (*Simple Sequence Repeat* – SSR, *Short Tandem Repeat* – STR) besteht aus einer kurzen Abfolge von tandemartigen Wiederholungen einer einfachen Sequenz, die maximal 12 bp lang ist. Sie liegt im Allgemeinen in nicht-kodierenden Bereichen und ist in hoher Dichte über das gesamte Genom verteilt (2% des Genoms). Am häufigsten ist die Wiederholung eines Motivs bestehend aus zwei Nukleotiden (Dinukleotidrepeat), seltener die Klasse der Trinukleotid- und Tetranukleotidwiederholungen. Die Mikrosatelliten-DNA ist hoch polymorph, das heißt, die Länge der Allele variiert in Abhängigkeit von der Anzahl der Sequenzwiederholungen in einem hohen Grad in der Bevölkerung. Die Allele der Mikrosatelliten werden nach den Mendelschen Regeln kodominant vererbt und lassen so einen Rückschluss über die Vererbung in einem Stammbaum zu. Dies macht sie zu geeigneten Markern für eine Kopplungsanalyse.

#### Auswahl der Mikrosatelliten-Marker und Design spezifischer Primer

Bei der Auswahl der Mikrosatelliten-Marker ist darauf zu achten, dass diese einen hohen Grad an Heterozygotie aufweisen, d.h., dass sie in einem großen Teil der Bevölkerung in zwei Allelen vorliegen. Die Heterozygotie wird als Wert von null bis eins angegeben. Der Wert Eins entspricht einer Heterozygotie von 100%. Ein hoher Heterozygotiegrad ist für die Effizienz der Kopplungsanalyse wichtig, damit in dem untersuchten Stammbaum möglichst viele informative Meiosen erhalten werden. Ist ein Mikrosatellit homozygot, d.h. das väterliche und/oder das mütterliche Allel weist dieselbe Anzahl an Sequenzwiederholungen auf, so lässt dies keinen Rückschluss zu, welches Allel weitervererbt wurde. Homozygote Loci sind somit nicht informativ. Weisen beide Eltern denselben heterozygoten Genotyp auf, ist die Hälfte der Meiosen nicht informativ [3]. In der vorliegenden Arbeit wurden nur Mikrosatelliten-Marker mit einer Heterozygotie größer als 0,70 ver-

---

<sup>1</sup><http://haplopainter.sourceforge.net>

wendet.

Für jede der acht möglichen Kopplungsregionen wurden zwei oder drei Mikrosatelliten, bevorzugt Tri- oder Tetranukleotidwiederholungen, ausgewählt. Alle in dieser Arbeit verwendeten Marker wurden der “Marshfield humanen genetischen Karte” der NCBI Map-Viewer Datenbank<sup>2</sup> entnommen. Die Sequenzen und Informationsangaben wurden mit der UCSC Genome Bioinformatics Datenbank<sup>3</sup> abgeglichen. Da für den Bereich auf Chromosom 8 nur ein dotierter Marker den Kriterien entsprach, wurde in der Genomsequenz des Assembly hg18<sup>3</sup> nach zusätzlicher Mikrosatelliten-DNA gesucht, die noch nicht benannt und näher untersucht worden ist. Hierfür wurde die Option “Repeating Elements by Repeat Masker” genutzt, die alle Wiederholungselemente und -marker in einer chromosomalen Region anzeigt. Es wurde der Dinukleotidrepeat  $21 \times \text{TG}$  in die Genotypisierung miteinbezogen. Alle ausgewählten Mikrosatelliten lagen innerhalb der Region der “folgenden Marker”, außer der Marker GATA43A04 auf Chromosom 1, AFM143XD8 und  $21 \times \text{TG}$  auf Chromosom 8, AFMA053XB5 auf Chromosom 7 und UT705 auf Chromosom 19. Diese lagen jeweils zwischen dem flankierenden und dem “folgenden Marker”. Sie sollten einer möglichen weiteren Eingrenzung der Kandidatenregion dienen.

Für die Detektion eines Mikrosatelliten benötigt man Oligonukleotide, die paarweise als Primer in der Polymerase-Kettenreaktion zur Amplifizierung von Mikrosatelliten-DNA eingesetzt werden. Die Oligonukleotide wurden mithilfe des Programms Primer3Plus<sup>4</sup> konstruiert. Die Primersequenzen wurden so gewählt, dass sie knapp außerhalb der Mikrosatelliten-DNA binden und das entstehende Produkt zwischen 100 und 500 bp lang ist. Dies bewirkt, dass die Länge des PCR-Produkts direkt von der Anzahl der Mikrosatelliten-Sequenzwiederholungen abhängig ist. Dabei ist darauf zu achten, dass der Primer nur an die eine ausgewählte Stelle im Genom bindet und diese nicht innerhalb eines Wiederholungselements liegt. Ferner wird eine feste Bindung nur erreicht, wenn keine SNPs oder andere Variationen in der Sequenz sind. Die Primer selbst sollten untereinander keine Primer-Dimere, Haarnadelstrukturen oder andere sekundäre Strukturen ausbilden. Überprüft wurden die Oligonukleotide mit einer in silico PCR und BLAT auf der UCSC Genome Bioinformatics Datenbank. An jeden Forward-Primer wurde die M13-Sequenz 5'-TGT AAA ACG ACG GCC AGT-3' angehängt.

Alle in dieser Arbeit verwendeten Primer wurden über Metabion International AG, Martinsried/Deutschland bezogen. Nach Lieferung wurde durch Zugabe von ultrasterilem  $\text{H}_2\text{O}$  eine Stocklösung der Konzentration 100 pMol/ $\mu\text{l}$  hergestellt. Durch eine 1 : 100-Verdünnung dieser Stocklösung wurde eine Primerkonzentration von 1 pMol/ $\mu\text{l}$  erreicht, die für die PCR der Mikrosatelliten-Marker benötigt wurde.

In der Tabelle 1 des Anhangs sind alle verwendeten Mikrosatelliten-Marker mit der Sequenz ihrer Primer, der Heterozygotität, der mittleren Produktgröße, der Wiederholungsart, der physikalischen Position, der PCR-Bedingung und der Annealingtemperatur  $T_A$  aufgeführt.

### PCR zur Generierung Fluoreszenz-markierter DNA-Fragmente

Eine Genotypisierung mit Hilfe von Mikrosatelliten-Markern wird erreicht durch eine Polymerase-Kettenreaktion (*polymerase chain reaction*, PCR) mit spezifischen Oligonukleotid-Primern, um eine ausreichende Menge an DNA für die Untersuchung zur Verfü-

<sup>2</sup><http://www.ncbi.nlm.nih.gov/sites/entrez?db=unists>

<sup>3</sup><http://genome.ucsc.edu>

<sup>4</sup><http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>

gung zu haben. Mithilfe der PCR wird der gewählte Nukleinsäureabschnitt, der zwischen dem Forward- und dem Reverse-Primer liegt, die für jeden Mikrosatelliten-Marker spezifisch sind, *in vitro* exponentiell amplifiziert.

Eine Standard-PCR besteht aus drei aufeinanderfolgenden Schritten: die Denaturierung, das Annealing und die Elongation. Die Denaturierung erfolgt bei 94°C und dient der Trennung der beiden Stränge der Ziel-DNA. Im nächsten Schritt wird die Temperatur auf die Annealingtemperatur  $T_A$  gesenkt.  $T_A$  richtet sich v.a. nach der Schmelztemperatur  $T_M$  der Primer und kann unter Anwendung verschiedener Formeln berechnet werden. Bei dieser hybridisieren die Oligonukleotid-Primer an die einzelsträngige Zielsequenz. Die darauf folgende Elongation erfolgt in der Regel bei 72°C, dem Temperaturoptimum der Taq-Polymerase. Diese verlängert die Primer an ihrem 5'-Ende und erzeugt wieder eine doppelsträngige DNA, die der ursprünglichen Template-DNA exakt gleicht. Dieser Zyklus wird nun mehrfach wiederholt und führt zu einer exponentiellen Steigerung der Template-DNA-Menge. Zusätzlich setzt man an den Anfang einen Schritt der initialen Denaturierung bei 94°C und als letzten Schritt eine finale Elongation bei 72°C, um die Arbeit der Polymerase zu beenden.

Um nach der PCR die Länge des jeweiligen PCR-Produkts mittels Elektrophorese und einem Laser-Detektionssystem erfassen zu können, muss einer der verwendeten Primer einen angehängten fluoreszierenden Farbstoff tragen. In der vorliegenden Arbeit wurde für das Markieren der PCR-Fragmente mit einem fluoreszierenden Farbstoff ein System mit drei Primern nach Schuelke (2000) [40] angewandt (vgl. Abbildung 4).

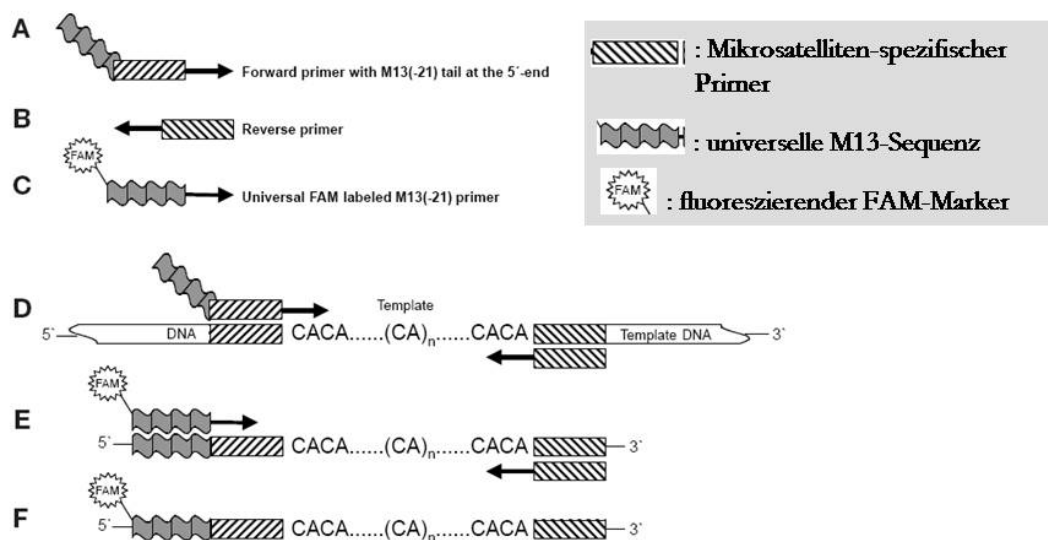


Abbildung 4: **Amplifikationsschema der PCR-Methode nach Schuelke (2000) [40].**

Man benötigte hierfür einen Sequenz-spezifischen Forward-Primer mit einem M13-Anhang an seinem 5'-Ende (A), einen Sequenz-spezifischen Reverse-Primer (B) und den universellen Fluoreszenz-markierten M13-Primer (C). Der Forward-Primer wurde in vierfach geringerer Menge wie der Reverse-Primer eingesetzt. Die Bedingungen für den Thermocycler (vgl. unten) sollten so gewählt werden, dass in den ersten 30 Zyklen der Forward-Primer mit seinem M13-Anhang in das PCR-Fragment eingebaut wird und allmählich verbraucht wird (D). Die Annealingtemperatur  $T_A$  lag zwischen 56°C und 61°C. In einem zweiten Schritt band an dieses Produkt der mit FAM-markierte universelle M13-Primer mit der Sequenz FAM-TGT AAA ACG ACG GCC AGT-3' (E). Da die Annealingtemperatur

peratur des universellen M13-Primers bei 53°C liegt, wurde diese in den nachfolgenden 8 Zyklen erniedrigt. So übernahm der FAM-M13-Primer die Rolle des Forward-Primers und fügte den Fluoreszenz-markierten Farbstoff in das PCR-Produkt mit ein. Das am Ende entstandene DNA-Fragment war Fluoreszenz-markiert und konnte mit einem Detektionssystem analysiert werden (**F**).

Für den Einsatz in der PCR wurde die genomische DNA der elf untersuchten Familienmitglieder mit ultrasterilem H<sub>2</sub>O auf eine Konzentration von 40 ng/µl verdünnt.

Verschiedene Stringenzbedingungen (- bzw. +) wurden unter Zugabe von DMSO erreicht. Zusätzlich wurde für jedes Primerpaar ein Temperaturgradient  $\Delta T$  von 10°C für die Annealingtemperatur T<sub>A</sub> ausgetestet, um die optimalen Bedingungen zu finden.

Die Ansätze für die PCR hatten folgende Zusammensetzung:

Reagenz	—	+
genomische DNA (40 ng/µl)	2 µl	2 µl
M13-Forward-Primer (1 pMol/µl)	1 µl	1 µl
Reverse-Primer (1 pMol/µl)	4 µl	4 µl
FAM-M13-Primer (1 pMol/µl)	4 µl	4 µl
dNTP-Mix (12,5 mM)	2 µl	2 µl
10× PCR Puffer (15 mM MgCl <sub>2</sub> )	2,5 µl	2,5 µl
h-Taq	0,5 µl	0,5 µl
DMSO	—	1 µl
ultrasteriles H <sub>2</sub> O	9 µl	8 µl
<b>Gesamt</b>	25 µl	25 µl

Falls mit dem - bzw. + Ansatz kein ausreichend gutes Ergebnis erzielt werden konnte, wurde ein Ansatz mit Q-Solution verwendet. Dieses Reagenz erleichtert die Amplifikation schwieriger Templates, indem es das Schmelzverhalten der DNA modifiziert [41].

Die Ansätze für die PCR mit Q-Solution enthielten folgende Substanzen:

Reagenz	
genomische DNA (40 ng/µl)	2 µl
M13-Forward-Primer (1 pMol/µl)	1 µl
Reverse-Primer (1 pMol/µl)	4 µl
FAM-M13-Primer (1 pMol/µl)	4 µl
dNTP-Mix (12,5 mM)	2 µl
10× PCR Puffer (15 mM MgCl <sub>2</sub> )	2,5 µl
h-Taq	0,5 µl
Q-Solution	2,5 µl
ultrasteriles H <sub>2</sub> O	6,5 µl
<b>Gesamt</b>	25 µl

Mit folgendem PCR-Programm wurden die Mikrosatelliten-Sequenzen amplifiziert:



Initiale Denaturierung	94°C	3 min	
Denaturierung	94°C	30 sec	
Annealing	T <sub>A</sub>	30 sec	30 Zyklen
Elongation	72°C	30 sec	
Denaturierung	94°C	30 sec	
Annealing	53°C	30 sec	8 Zyklen
Elongation	72°C	30 sec	
Schlusselongation	72°C	10 min	
	10°C	bis zur weiteren Bearbeitung	

### Agarosegel-Elektrophorese

Die Agarosegel-Elektrophorese ist eine effektive Methode, um DNA-Fragmente aufzutrennen und die Effizienz, Spezifität und Korrektheit der PCR zu überprüfen. Agarose wurde in einer Konzentration von 1,5% in 1× TBE Elektrophoresepuffer aufgekocht und gelöst. Mit Hilfe eines Gelschlittens und Gelkämmen wurde ein Gel mit Taschen gegossen. Um DNA in dem Agarosegel sichtbar zu machen, wurde in die aufgekochte Lösung ein Tropfen Ethidiumbromid-Lösung (0,07%) gemischt. Das feste Gel wurde in eine Elektrophoresekammer gelegt und diese mit 1× TBE aufgefüllt. Nachdem je 5 µl PCR-Produkt mit 2 µl Ladepuffer (*loading buffer*) Xylencyanolblau vermischt in die Taschen pipetiert worden war, wurde eine Spannung von 160 Volt angelegt. Dies bewirkte, dass die negativ geladenen DNA-Fragmente in Abhängigkeit ihrer Größe von der Kathode zur positiv geladenen Anode wanderten. Je kleiner die DNA-Fragmente waren, desto schneller durchliefen sie das Agarosegel, dessen Poren kleinere Fragmente weniger zurückhielten. Nach etwa 20 Minuten war die DNA weit genug gelaufen und das Ergebnis konnte unter einem UV-Schirm betrachtet werden. Anhand des mitgelaufenen DNA-Größenmarkers, dem 1 Kb Plus DNA-Leiter, der DNA-Fragmente in regelmäßigen Abständen enthielt, konnte die Länge der DNA-Fragmente in den Proben abgeschätzt werden.

Für die weitere Verwendung wurden die PCR-Produkte 1 : 8 mit ultrasterilem H<sub>2</sub>O verdünnt.

### Fragmentanalyse mittels Kapillarelektrophorese und anschließender Größenbestimmung

Die Fragmentanalyse diente der Längenbestimmung der Fluoreszenz-markierten DNA-Fragmente. Sie wurde mit dem ABI3130xl Genetic Analyser durchgeführt und mit der Software GeneMapper Version 3.7 (Applied Biosystem) ausgewertet. Die Elektrophorese des mit FAM-markierten PCR-Produkts erfolgte zusammen mit einem internen Längenstandard, dem Rox-Längenstandard (Applied Biosystem).

Der Ansatz für den Fragmentanalyselauf enthielt:

Reagenz	
verdünnte PCR-Probe (1 : 8)	1 µl
GeneScan <sup>TM</sup> 500 Rox <sup>TM</sup> Size Standard	0,25 µl
HiDi <sup>TM</sup> Formamide	11 µl
<b>Gesamt</b>	<b>12,25 µl</b>

Der Längenstandard kompensierte Abweichungen von Lauf zu Lauf und von Kapillare zu Kapillare und ermöglichte damit den Vergleich von DNA-Fragmenten nach ihrer Länge. Um die Längen der Fragmente eindeutig bestimmen zu können, war eine komplette Denaturierung der PCR-Fragmente und der Fragmente des Längenstandards nötig. Die Ansätze für den Fragmentanalyselauf wurden drei Minuten bei 94°C im Thermocycler denaturiert. Danach konnten die Ansätze in die Mikrotiterplatten für den ABI3130xl Genetic Analyser überführt werden.

Das Fragmentanalysegerät arbeitet mit Kapillaren, die einen Innendurchmesser von 50 µm haben und mit einem Polymer gefüllt sind. Die Kapillare taucht mit dem einen Ende, an dem sich die Kathode befindet, in die vorgelegte Probe ein. Durch Anlegen einer Spannung kommt es zu einem Stromfluss zwischen der Kathode (-) und der Anode (+), die sich am anderen Ende der Kapillare befindet. Dieser bewirkt, dass negativ geladene Teilchen und damit DNA-Fragmente aus der Probe aufgenommen werden. Nach diesem Vorgang der elektrokinetischen Injektion wird die Kathode für die Elektrophorese in einem Puffer platziert, damit keine weiteren DNA-Fragmente aufgenommen werden. In dem angelegten elektrischen Feld wandern geladene Teilchen unterschiedlich schnell, je nach der Größe der Fragmente, durch das Polymer der Kapillare hin zur Anode. Am Ende der Kapillare passieren die Fragmente ein Detektionsfenster, das im Strahl eines Argonlasers liegt. Dieser Laser regt die fluoreszenzmarkierten PCR-Produkte zur Emission an. So können die emittierten Fluoreszenzsignale von einer digitalen Kamera erfasst werden [42].

Nach Ende der Kapillarelektrophorese wurden die erhaltenen Rohdaten durch die GeneMapper®Software Version 3.7 analysiert. Die Längen der PCR-Produkte stellten sich als Fluoreszenzsignale oder Peaks in einem Koordinatensystem dar. Auf der X-Achse wurden die Produktlängen (in Basen), auf der Y-Achse die Intensität der fluoreszierenden Signale aufgetragen. Anhand des verwendeten Größenstandards, dessen Fragmentlängen bekannt sind, berechnete die GeneMapper®Software die Länge der fluoreszenzmarkierten PCR-Produkte. Der Rox-Standard wurde in der graphischen Darstellung der Rohdaten als rote Linie sichtbar. Die verschiedenen Peaks des Graphen wurden in dem Größenbereich ausgewertet, der für den Mikrosatelliten-Marker bekannt ist. War ein Individuum für den untersuchten Mikrosatelliten-Marker homozygot, das heißt, es trug zwei Allele gleicher Länge, war nur ein Peak zu sehen. Bei Heterozygotie ergaben sich zwei Peaks pro Marker. Anschließend wurde geprüft, welcher Elternteil welches Allel an seine Nachkommen weitergegeben hat.

Nachfolgend ist jeweils ein Beispiel für ein heterozygotes (informatives) und homozygotes (nicht informatives) Signal in einer Familiensituation abgebildet.

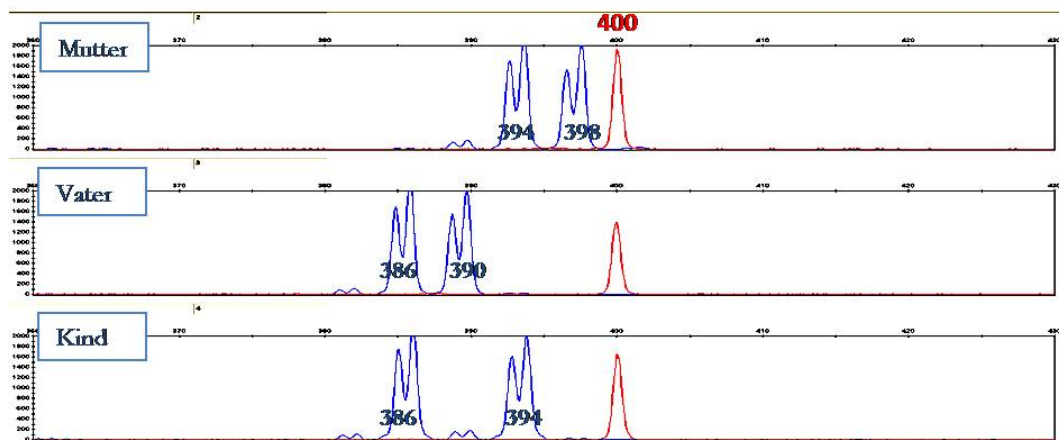


Abbildung 5: **Heterozygote Familiensituation.**

Die Signale der FAM-markierten Fragmentlängen sind als blaue Linie, der Rox-Standard als rote Linie in der graphischen Darstellung der Rohdaten dargestellt. Mutter und Vater sind in diesem Beispiel jeweils heterozygot für den analysierten Mikrosatelliten, einen Tetranukleotidrepeat. Es kann nachvollzogen werden, welches Allel das Kind von welchem seiner Eltern geerbt hat.

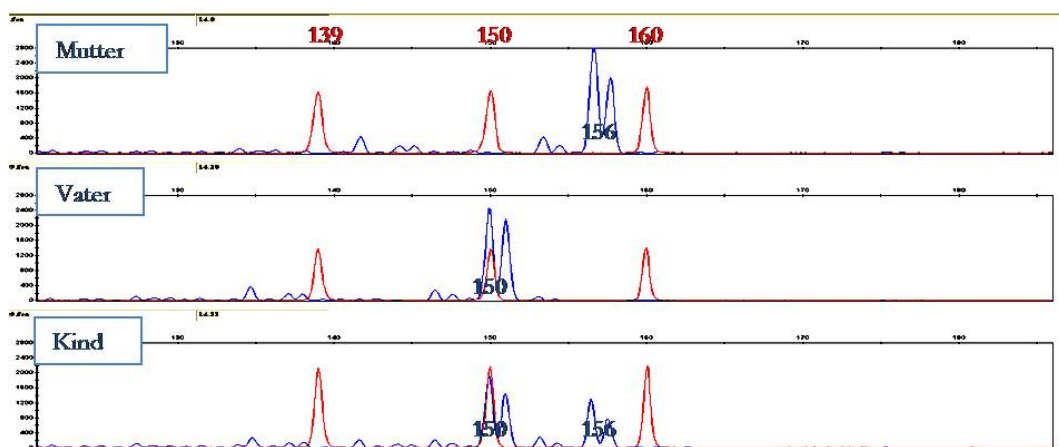


Abbildung 6: **Homozygote Familiensituation.**

Die Signale der FAM-markierten Fragmentlängen sind als blaue Linie, der Rox-Standard als rote Linie in der graphischen Darstellung der Rohdaten dargestellt. Mutter und Vater sind in diesem Beispiel jeweils homozygot für den analysierten Mikrosatelliten, einen Trinukleotidrepeat. Es kann nicht nachvollzogen werden, welches der beiden elterlichen Allele das Kind von Vater bzw. Mutter geerbt hat.

Bei der Interpretation der Rohdaten war darauf zu achten, dass durch die Syntheseungenauigkeit der Taq-Polymerase bei der PCR für jedes Allel zusätzliche Bande, sogenannte Stotterbanden (*stutter* - oder *shadow bands*), entstanden. Ursächlich dafür ist, dass Wiederholungseinheiten innerhalb der Mikrosatelliten-DNA von der Polymerase übersprungen werden und so in der PCR zum ursprünglichen Produkt vervielfältigt werden.

Dinukleotidmarker neigen besonders zu Replikationsverschiebungen und sind wesentlich schwerer auszuwerten als Trinukleotid- und Tetranukleotidmarker (vgl. Abbildung 5: Tetranukleotidrepeat, Abbildung 6: Trinukleotidrepeat und Abbildung 7: Dinukleotidrepeat).

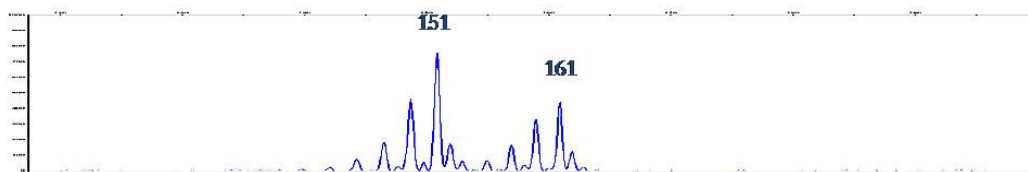


Abbildung 7: **Stotterbanden eines heterozygoten Dinukleotidrepeats.**

*Die Person ist heterozygot für den analysierten Dinukleotidrepeat. Ein Allel ist 151 Basen, das andere 161 Basen lang. Durch die Synthesegenauigkeit der Taq-Polymerase entstehen für jedes Allel sog. Stotterbanden, die sich als niedrigere Peaks vor allem vor dem Hauptpeak darstellen und zu Überlagerungen führen können.*

## 2.4 Analyse der Kandidatengene

Das "Human Genome Project" hat das gesamte menschliche Genom sequenziert und genetische und physikalische Karten des Genoms zur Verfügung gestellt [43]. Unter Anwendung dieser Karten, die in öffentlichen Datenbanken zugänglich sind, lässt sich herausfinden, welche Gene in einer bestimmten genomischen Region vorliegen. In dieser Arbeit wurde mit Hilfe der UCSC Genome Bioinformatics Datenbank in der Karte des humanen Genoms hg18 vom März 2006 nach den Genen gesucht, die in den identifizierten Kopplungsregionen liegen. Diese wurden nach ihrer Funktion und möglichen Rolle als Kandidatengene gelistet. Kandidatengene stellten solche Gene dar, die in einer Kopplungsregion liegen und deren Funktion oder vielmehr deren Defekt mit dem Phänotyp der Erkrankung Brustkrebs in Zusammenhang gebracht werden kann. Die Kandidatengenenanalyse erfolgte mithilfe der Recherche der Genfunktion auf verschiedenen Datenbanken, v.a. der GeneCards®<sup>5</sup>, der PubMed<sup>6</sup> und der UCSC Genome Bioinformatics Datenbank.

## 2.5 Sequenzierung von Kandidatengen

Die Sequenzierung ist eine Methode, die der Aufklärung der Abfolge der Basen in DNA-Fragmenten dient und so ermöglicht, Abweichungen in einem Sequenzvergleich zu finden.

### 2.5.1 Das Prinzip der DNA-Sequenzierung

Die Sequenzierung nach Sanger (1977) [44] beruht auf dem Prinzip der Kettenabbruchmethode. Der Kern dieser Methode ist die Synthese eines DNA-Stranges, der zu einer einzelsträngigen Matrize komplementär ist. Im Verlauf der Sequenz-Reaktion baut eine DNA-Polymerase an das freie 3'-OH Ende eines Oligonukleotid-Primers zur Ziel-DNA komplementäre Basen ein. Als Bausteine werden sowohl Desoxynukleotide (dNTPs) als auch Didesoxynukleotide (ddNTPs), die im Vergleich mit den dNTPs zusätzlich zu dem 2'-C Atom auch am 3'-C Atom der Ribose keine OH-Gruppe tragen, verwendet. Der Abbruch der Synthese des DNA-Stranges erfolgt sequenzspezifisch durch Einbau eines ddNTPs, das von der DNA-Polymerase nicht mehr verlängert werden kann, da eine

<sup>5</sup><http://www.genecards.org>

<sup>6</sup><http://www.ncbi.nlm.nih.gov/pubmed>

Phosphodiesterbindung zum nächsten Nukleotid nicht gebildet werden kann. In dem Reaktionsgemisch für die Sequenzierungs-PCR befinden sich dNTPs und ddNTPs in einem bestimmten Verhältnis zueinander, das gewährleisten soll, dass an jeder Basenposition eine später detektierbare Fraktion der Moleküle durch den Einbau eines ddNTPs terminiert wird, während die größere Fraktion der Moleküle durch den Einbau eines dNTPs verlängert wird. Die Abbruchmoleküle ddATP, ddGTP, ddCTP und ddTTP sind mit einem für die entsprechende Base spezifischen Fluoreszenzfarbstoff markiert (ddATP = grün, ddGTP = schwarz, ddCTP = blau und ddTTP = rot). Dies ermöglicht in der Auswertung die vier verschiedenen Basen aufgrund unterschiedlicher spektraler Eigenschaften dieser Farbstoffe zu unterscheiden. In der Sequenzierungsreaktion entsteht so ein heterogenes PCR-Produkt aus DNA-Fragmenten unterschiedlicher Länge, dessen zuletzt eingebautes Didesoxynukleotid fluoreszenzmarkiert ist. Je später beim Synthetisieren des komplementären DNA-Stranges ein ddNTP eingebaut wird, desto länger wird das Fragment. Die Längendifferenz zwischen den Produkten beträgt jeweils ein Nukleotid.

Die Reaktionsprodukte werden anschließend automatisch im ABI3130xl Genetic Analyser mittels Kapillarelektrophorese über ein Polymer aufgetrennt. Die negativ geladenen DNA-Fragmente wandern abhängig von ihrer Größe unterschiedlich schnell in Richtung der positiv geladenen Anode und werden am Ende der Kapillare von einem Laser erfasst. Voraussetzung für den Erfolg der Auftrennung ist, dass die DNA-Fragmente denaturiert sind und während der Elektrophorese auch einzelsträngig bleiben. Bei partieller oder vollständiger Basenpaarung innerhalb des Stranges würden sich die Elektrophorese-Eigenschaften im Vergleich zum linearen Strang verändern und keine Korrelation zwischen Wanderungszeit und Sequenzlänge erlauben. Das Laser-Licht regt die fluoreszierenden Farbstoffe zur Emission an, deren Strahlung von einem Detektor erfasst wird. Der Detektor registriert die verschiedenen Fluoreszenz-markierten Enden der DNA-Stücke in der zeitlichen Reihenfolge wie sie am Laser vorbeilaufen. Eine angeschlossene Software interpretiert diese Signale und rechnet aus der Wellenlänge der detektierten Signale auf die jeweils eingebaute Base zurück. Dadurch erhält man die Basensequenz, deren Abfolge zur ursprünglich sequenzierten Template-DNA komplementär ist.

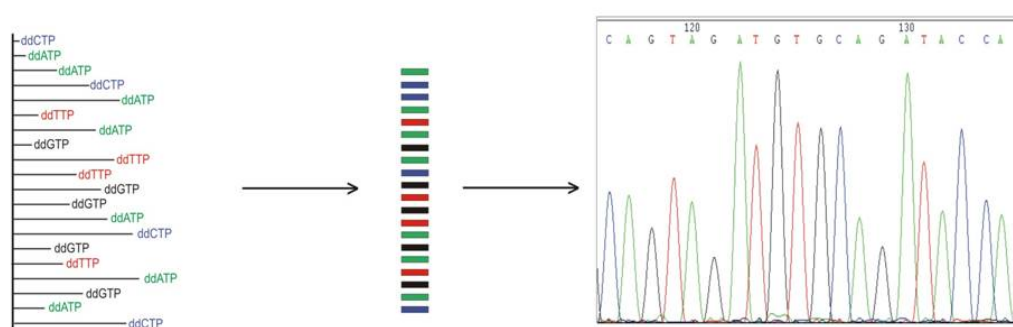


Abbildung 8: **Ablauf der Sequenzierung.**

In der Sequenzier-Reaktion entstehen durch Einbau der Fluoreszenz-markierten ddNTPs verschieden lange DNA-Fragmente. Die Auftrennung der DNA-Fragmente erfolgt mittels Kapillarelektrophorese nach ihrer Länge; kurze Fragmente wandern schneller durch das Polymer als lange Fragmente. Das Computerprogramm übersetzt die vom Detektor registrierten Lichtsignale in die Basensequenz.

### 2.5.2 Durchführung der DNA-Sequenzierung

#### PCR zur Amplifizierung der Ziel-DNA

Zunächst wurde die DNA-Region, die später sequenziert werden sollte, mittels einer PCR amplifiziert. Die Oligonukleotidprimer für diese Reaktion wurden mithilfe des Programms Primer3Plus konzipiert (vgl. Abschnitt 2.3.3). Dabei wurde darauf geachtet, dass diese mindestens 35 Basen vor bzw. hinter dem zu sequenzierenden Exon binden, da die Anfangsbereiche der Sequenzen häufig unvollständig auszuwerten sind. Ferner sollten auch Mutationen in den Spleiß-Konsensus-Sequenzen im Intron vor und nach der kodierenden Sequenz untersucht werden. Weiter war darauf zu achten, dass der Abstand beider Primer die maximal lesbare Sequenzlänge von 600 bp nicht überschreitet. Bei größeren Exonen wurden überlappende Primerpaare gewählt.

Im Anhang in Tabelle 2 sind alle sequenzierten Exone mit der Sequenz ihrer Primer, der Produktgröße, der PCR-Bedingung und der Annealingtemperatur  $T_A$ , sowie die für das jeweilige Exon untersuchten Individuen aufgeführt.

Nach Lieferung dieser Primer von International AG, Martinsried/Deutschland wurde durch Zugabe von ultrasterilem  $H_2O$  eine Stocklösung der Konzentration 100 pMol/ $\mu$ l hergestellt. Durch eine 1 : 10 Verdünnung dieser Stocklösung erhielt man eine Primerkonzentration von 10 pMol/ $\mu$ l, die für die PCR benötigt wurde. Die Zusammensetzung des PCR-Ansatzes unterschied sich von den Ansätzen für die PCR zur Genotypisierung der Mikrosatelliten-Marker.

Die Ansätze für die PCR hatten folgende Zusammensetzung:

Reagenz	—	+
genomische DNA (40 ng/ $\mu$ l)	2 $\mu$ l	2 $\mu$ l
Forward-Primer (10 pMol/ $\mu$ l)	1 $\mu$ l	1 $\mu$ l
Reverse-Primer (10 pMol/ $\mu$ l)	1 $\mu$ l	1 $\mu$ l
dNTP-Mix (12,5 mM)	2 $\mu$ l	2 $\mu$ l
10× PCR Puffer (15 mM $MgCl_2$ )	2,5 $\mu$ l	2,5 $\mu$ l
h-Taq	0,5 $\mu$ l	0,5 $\mu$ l
DMSO	—	1 $\mu$ l
ultrasteriles $H_2O$	16 $\mu$ l	15 $\mu$ l
<b>Gesamt</b>	25 $\mu$ l	25 $\mu$ l

Verschiedene Stringenzbedingungen (- bzw. +) wurden unter Zugabe von DMSO erreicht. Zusätzlich wurde für jedes Primerpaar ein Temperaturgradient von 10°C für die Annealingtemperatur  $T_A$  ausgetestet, um die optimalen Bedingungen für die PCR zu finden.

Wenn damit kein ausreichend gutes Ergebnis erzielt werden konnte, wurde ein Ansatz mit Q-Solution verwendet.

Die Ansätze für die PCR mit Q-Solution enthielten folgende Substanzen:

Reagenz	
genomische DNA (40 ng/ $\mu$ l)	2 $\mu$ l
Forward-Primer (10 pMol/ $\mu$ l)	1 $\mu$ l
Reverse-Primer (10 pMol/ $\mu$ l)	1 $\mu$ l
dNTP-Mix (12,5 mM)	2 $\mu$ l
10 $\times$ PCR Puffer (15 mM MgCl <sub>2</sub> )	2,5 $\mu$ l
h-Taq	0,5 $\mu$ l
Q-Solution	2,5 $\mu$ l
ultrasteriles H <sub>2</sub> O	13,5 $\mu$ l
<b>Gesamt</b>	25 $\mu$ l

Für einige besonders schwierig zu amplifizierende DNA-Regionen wurde statt der gewöhnlichen h-Taq die AccuPrime<sup>TM</sup> GC-Rich DNA Polymerase (invitrogen<sup>TM</sup>) verwendet. Der zugehörige 5 $\times$  Reaktionspuffer A beinhaltet dNTPs, MgSO<sub>4</sub> und temperaturstabile AccuPrime<sup>TM</sup> Zusatzproteine.

Die Ansätze für die PCR mit AccuPrime<sup>TM</sup> GC-Rich DNA Polymerase enthielten:

Reagenz	
genomische DNA (40 ng/ $\mu$ l)	2 $\mu$ l
Forward-Primer (10 pMol/ $\mu$ l)	1 $\mu$ l
Reverse-Primer (10 pMol/ $\mu$ l)	1 $\mu$ l
5 $\times$ Puffer A	5 $\mu$ l
AccuPrime <sup>TM</sup> GC-Rich DNA Polymerase	0,5 $\mu$ l
ultrasteriles H <sub>2</sub> O	15,5 $\mu$ l
<b>Gesamt</b>	25 $\mu$ l

Mit folgendem Touch-Down PCR-Programm wurde die DNA vor der Sequenzier-Reaktion amplifiziert:

Initiale Denaturierung	94°C	2 min	
Denaturierung	94°C	30 sec	6 Zyklen mit +dT = -1 für T <sub>A</sub>
Annealing	T <sub>A</sub> + 6°C	30 sec	
Elongation	72°C	30 sec	
Denaturierung	94°C	30 sec	30 Zyklen
Annealing	T <sub>A</sub>	30 sec	
Elongation	72°C	30 sec	
Schlusselongation	72°C	5 min	
	10°C	bis zur weiteren Bearbeitung	

Nachdem für jedes Primerpaar die optimale Annealingtemperatur und Pufferzusammensetzung ermittelt worden waren, wurden mit diesen PCRs für zwei Trägerinnen des krankheitsspezifischen Haplotypen und für eine Kontrollperson (II-5, II-6 und III-5 bzw. II-12,

II-13 und III-9) durchgeführt.

### PCR-Produkt Quantifizierung

Nach Ablauf der PCR wurden jeweils 5 µl aus jedem Ansatz zusammen mit 2 µl Ladepuffer auf ein 1,5%-iges Agarosegel geladen. Nach einer halbstündigen Elektrophorese bei 160V wurde das Ergebnis unter dem UV-Schirm beurteilt (vgl. Abschnitt 2.3.3).

### Verdau

Bevor das amplifizierte PCR-Produkt in die Sequenzier-Reaktion eingesetzt werden konnte, musste dieses in einem nächsten Schritt für 30 Minuten bei 37°C verdaut werden. Zu 1 µl PCR-Produkt wurde Exonuklease 1 und Antarctic Phosphatase gegeben, die bewirkten, dass Primerreste, unspezifische PCR-Produkte und noch vorhandene dNTPs abgebaut wurden.

Ein Ansatz für den Verdau enthielt:

Reagenz	
PCR-Produkt	1 µl
Exonuklease 1	0,1 µl
Antarctic Phosphatase	0,25 µl
ultrasteriles H <sub>2</sub> O	3,65 µl
<b>Gesamt</b>	5 µl

Um bei der anschließenden Sequenzier-Reaktion von beiden komplementären DNA-Strängen die Basensequenz ermitteln zu können, wurden für jedes PCR-Produkt zwei Verdau-Reaktionen angesetzt. Nach dem Verdau wurden die Proben für 15 Minuten auf 80°C erhitzt. Dies inaktivierte die beiden Enzyme.

### Sequenzier-Reaktion

Für die Sequenzier-Reaktion wurden zu je 5 µl des verdauten PCR-Produkts 5× BigDye Sequenzierungs-Puffer, Big Dye® Terminator Reaktions-Mix und Forward- oder Reverse-Primer in einer Konzentration von 10 pMol/µl gegeben.

Die Ansätze setzten sich folgendermaßen zusammen:

Reagenz	
verdautes PCR-Produkt	5 µl
Forward- oder Reverse-Primer	1 µl
5× Puffer	2 µl
BigDye Reaktions-Mix	0,3 µl
ultrasteriles H <sub>2</sub> O	1,7 µl
<b>Gesamt</b>	10 µl

Das Programm der Sequenzier-Reaktion für den Thermocycler lautete wie folgt:



Initiale Denaturierung	94°C	5 min	
Denaturierung	94°C	30 sec	
Annealing	58°C	30 sec	25 Zyklen
Elongation	60°C	4 min	
Schlusselongation	60°C	5 min	
	15°C	bis zur weiteren Bearbeitung	

Im Vergleich zu einer PCR mit zwei Primern stieg die Anzahl der amplifizierten DNA-Fragmente bei der Sequenzier-Reaktion nur linear an.

### Fällung

Nach der Sequenzier-Reaktion mussten die Proben gefällt werden. In einem ersten Schritt wurde zu jeder Probe 3 µl 3M Natriumacetat (NaAc) und 50 µl 100%-iger Ethanol gegeben. Diese wurden anschließend für 20 Minuten bei 4300 rpm und 23°C zentrifugiert. Der Überstand wurde abgenommen und verworfen; die Proben wurden über Kopf auf einem Blatt Küchenkrepp mit Quick Run kurz bis etwa 800rpm zentrifugiert. In einem zweiten Schritt wurde auf das Pellet 60 µl 70%-iger Ethanol gegeben und die Probe auf dieselbe Weise gewaschen. Am Ende wurde das Pellet in 10 µl HiDi<sup>TM</sup> Formamide gelöst. Die so vorbereiteten Proben waren fertig für die Analyse mit dem ABI3130xl Genetic Analyser.

### Auswertung der Sequenzen

Die Rohdaten wurden mit den Programmen SeqMan<sup>TM</sup> II, Chromas Version 2.23 und ApE-A Version 1.12 ausgewertet (vgl. Abschnitt 2.1.1). Die erhaltenen Sequenzen wurden mit der Zielsequenz des hg18 vom März 2006 aus der öffentlichen UCSC Genome Bioinformatics Datenbank verglichen und auf Abweichungen hin untersucht. Homozygote Mutationen sind von heterozygoten, bei denen nur ein DNA-Strang betroffen ist, zu unterscheiden. Bei homozygoten Mutationen sind die komplementären DNA-Stränge gleichermaßen betroffen. Sie zeigen sich im Vergleich zur Zielsequenz als Basenaustausch (Punktmutation), als Einfügen einer zusätzlichen Base oder mehrerer zusätzlicher Basen (Insertion) oder als Fehlen einer oder mehrerer Basen (Deletion). Bei einer heterozygoten Punktmutation sind zwei kleinere überlagerte Peaks zu sehen, bei einer heterozygoten Insertion oder Deletion kommt es danach zu einer Überlagerung der Basen des nicht mutierten Stranges mit denen des mutierten Stranges.

## 3 ERGEBNISSE

### 3.1 Familiäre Kopplungsanalyse

Für elf Individuen der Brustkrebs-Familie (II-5, II-6, II-7, II-10, II-12, II-13, III-1, III-5, III-6, III-7 und III-9) wurde eine genomweite Genotypisierung vorgenommen. In die Auswertung der Genotypisierungsdaten für die Kopplungsanalyse mit Merlin wurden 35.642 SNPs einbezogen. Das Muster des familiären Risikos in der Familie zeigte an, dass ein mutantes Allel wahrscheinlich dominant wirkt. Deswegen wurde die parametrische Kopplungsanalyse unter Annahme eines dominanten Modells als die wichtigste Analyse betrachtet. Alternativ wurden auch Analysen mit einem rezessiven Modell und nicht-parametrische Analysen durchgeführt.

#### 3.1.1 Annahme 1: Alle an Brustkrebs erkrankte Familienmitglieder tragen den genetischen Defekt

In der parametrischen Analyse wiesen unter Annahme eines dominanten Vererbungsmodells mehrere chromosomale Regionen auf Kopplung hin. Der höchste multipoint LOD-Wert (mLOD) von 1,8 wurde für eine Region auf dem langen Arm des X-Chromosoms (Xq) erzielt. Auf den Autosomen erreichte ein Locus auf 17q mit einem mLOD von 1,35 den höchsten Wert. Drei weitere Regionen zeigten einen  $mLOD > 0,8$ : ein Locus auf Chromosom 19 ( $mLOD \sim 1,1$ ), einer auf 8p ( $mLOD \sim 0,94$ ) und ein weiterer auf 20q ( $mLOD \sim 0,87$ ). In der nicht-parametrischen Kopplungsanalyse zeigten die fünf beschriebenen Regionen jeweils die höchsten Werte für  $Z_{mean}$ .

Die Kopplungsregionen auf Xq und 17q wurden in der Haplotypenanalyse und der Analyse mit Mikrosatelliten-Markern genauer untersucht.

#### Kopplungsregion Xq

Die Haplotypenanalyse der Genotypisierungsdaten offenbarte für die Region mit einem Kopplungssignal auf Xq gemeinsame Haplotypen, die betroffene Frauen und Überträgerinnen der Familie gemeinsam haben (vgl. Abbildung 9). Die Individuen II-5, II-7, II-10, II-12, III-1, III-5, III-6, III-7 und III-9 tragen den rot gekennzeichneten krankheitsspezifischen Haplotypen. Die Genotypisierung von drei Mikrosatelliten-Markern in dieser Region bestätigte den Risikohaplotypen und zeigte die Allele auf, die mit der Krankheit segregieren. Für den Dinukleotidrepeat AFMB018WD9 ist das Allel mit 172 Basen, für den Tetranukleotidrepeat TATC043 das Allel mit 172 Basen und für den Dinukleotidrepeat AFMA065WD9 das Allel mit 166 Basen mit der Erkrankung assoziiert.

Die Kandidatenregion überspannt einen Bereich von 14,53 Mb in der Region Xq27.2-q28. Als flankierender zentromerischer Marker konnte SNP\_A-1687176 identifiziert werden. III-1 ist homozygot für das Allel 1, während II-10, III-5, III-6 und III-9 homozygot für das Allel 2 sind. Ab diesem SNP tragen die betroffenen Personen nicht mehr denselben Haplotypen und SNP\_A-1687176 stellt die Grenze der Kopplungsregion dar. Da telomerwärts keine weitere Rekombination stattgefunden hat, reicht die Kandidatenregion bis zum Xq-Telomer.

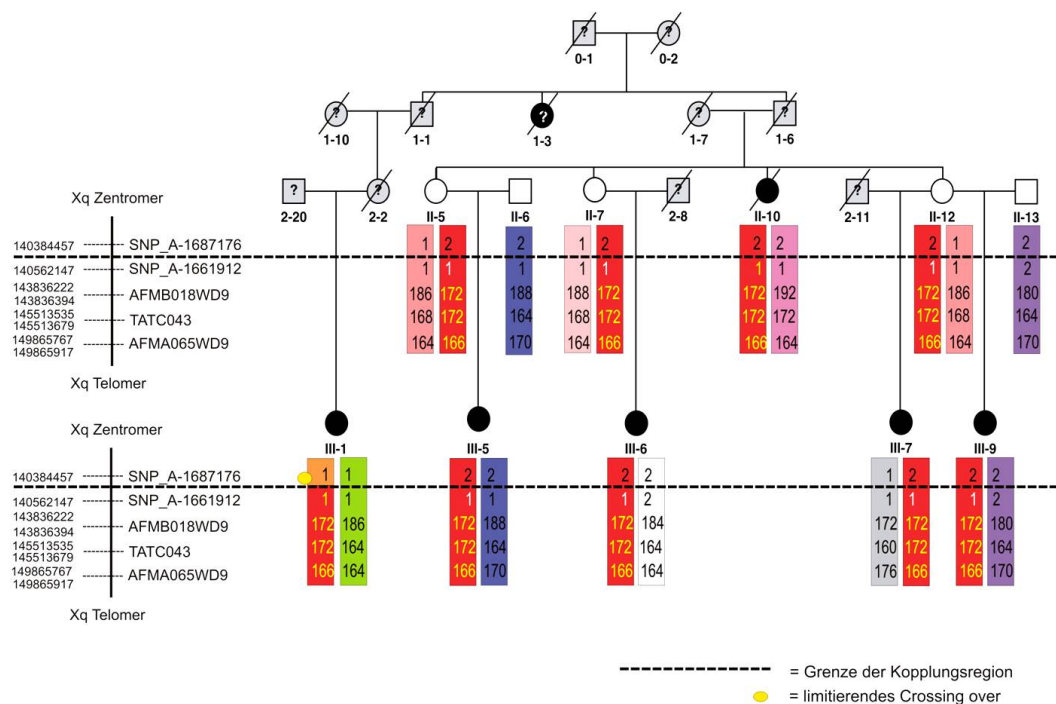


Abbildung 9: **Haplotypenanalyse des Chromosom X.**

Der Haplotyp in gelber Schrift und rotem Hintergrund segregiert mit der Erkrankung, die Allele in weißer Schrift gehören zu Markern, die nicht informativ waren. SNP\_A-1687176 ist nicht mehr mit der Erkrankung gekoppelt und stellt die Grenze der Kandidatenregion Xq dar.

### Kopplungsregion 17q

Die genauere Analyse der potenziellen Kandidatenregion auf 17q ergab, dass mit Ausnahme der Person III-1 alle erkrankten Frauen und obligaten Überträgerinnen den krankheitsassoziierten Haplotypen tragen (vgl. Abbildung 10). Person III-1 könnte somit eine Phänokopie darstellen. Sie trägt nicht den Risikohaplotypen. Für den krankheitsspezifischen Haplotypen wurden zwei Mikrosatelliten-Marker einer Fragmentlängenanalyse unterzogen. Das Allel mit 219 Basen für den Dinukleotidrepeat AFM338XG5 bzw. der Länge von 469 Basen für den Tetranukleotidrepeat GATA49C09 segregiert in II-5, II-7, II-10, II-12, III-5, III-6, III-7 und III-9 mit der Erkrankung.

Die Kandidatenregion überspannt einen Bereich von 24,08 Mb in der Region 17q21.33-q25.1. SNP\_A-1715757 ist der zentromerwärts flankierende Marker der Kopplungsregion. Das limitierende Crossing-over konnte zwischen diesem und dem folgenden Marker SNP\_A-1689534 nachvollzogen werden. In Richtung Telomer wird der mögliche Krankheitsloкус von SNP\_A-1656243 eingegrenzt. Ein Rekombinationsereignis hat zwischen SNP\_A-1656243 und SNP\_A-1674114 stattgefunden.

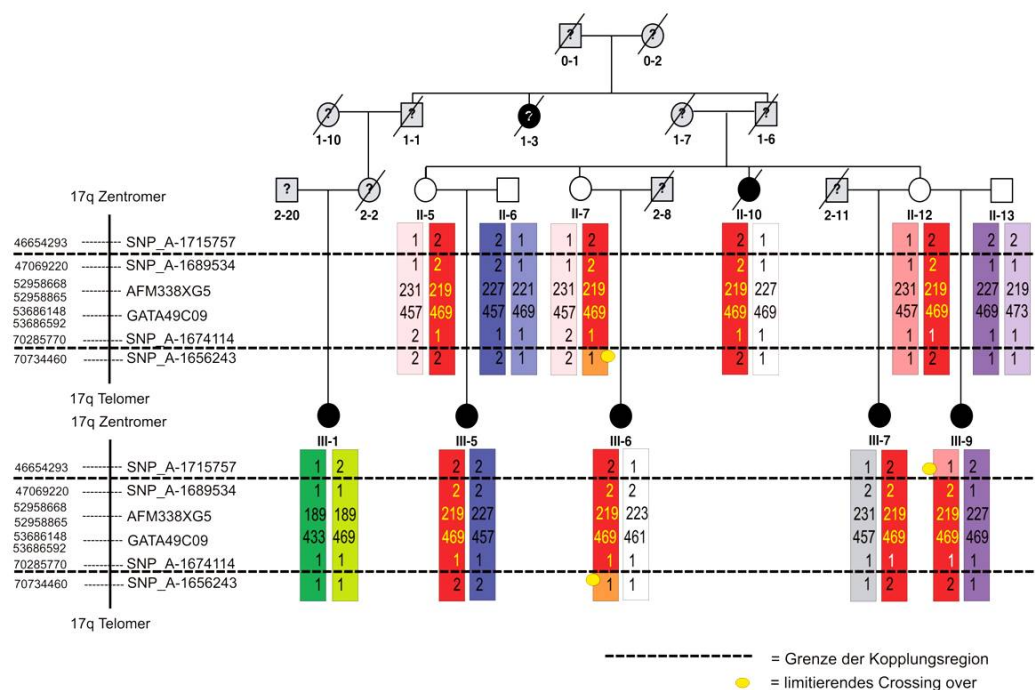


Abbildung 10: **Haplotypenanalyse des Chromosom 17.**

Der Haplotyp in gelber Schrift und rotem Hintergrund segregiert mit der Erkrankung, die Allele in weißer Schrift gehören zu Markern, die nicht informativ waren. *SNP\_A-1715757* und *SNP\_A-1656243* sind nicht mehr mit der Erkrankung gekoppelt und stellt die Grenze der Kandidatenregion 17q dar.

### 3.1.2 Annahme 2: Eine der betroffenen Personen im Stammbaum erkrankte an einem sporadischen Brustkrebs

Annahme 2 geht davon aus, dass eine betroffene Person nicht aufgrund des in der Familie mit der Erkrankung segregierenden genetischen Defekts, sondern aufgrund sporadischer Genese an Brustkrebs erkrankte. Sie stellt damit lediglich eine Phänokopie dar. In der Stammbaumditei wurde diese Person als fehlender Phänotyp ("unknown") angegeben. Bei allen anderen betroffenen Individuen wurde die Segregation der Erkrankung mit einem familiär bedingten genetischen Defekt angenommen. Die Berechnung erfolgte, indem nacheinander die Betroffenen III-1, III-5, III-6, III-7, III-9 und II-10 auf "unknown" gesetzt wurden. Die dargelegten Ergebnisse basieren in der parametrischen Kopplungsanalyse auf einem dominanten Vererbungsmodell.

Als mögliche Kandidatenregionen für eine Kopplung wurden bei der Analyse unter dieser Annahme solche chromosomalen Bereiche in Betracht gezogen, die in der parametrischen Analyse einen multipoint LOD-Wert  $> 1,4$  auf den Autosomen erzielten. Das X-Chromosom wurde aufgrund des vermuteten autosomal dominant vererbten Gendefekts in der Familie nicht explizit untersucht. Jedoch zeigte sich im Vergleich zu der identifizierten Kandidatenregion auf dem X-Chromosom unter Annahme 1 (vgl. Abschnitt 3.1.1), dass diese auch unter Annahme 2 erhöhte mLOD-Werte erreichte und jeweils die identische chromosomale Region mit den gleichen Grenzmarkern überspannte.

Die folgenden Abbildungen 11 bis 16 zeigen die graphische Darstellung der parametrischen Kopplungsanalyse unter Annahme 2.

Die mLOD-Werte sind gegen die Position in der genetischen Karte aufgetragen. Auf der Y-Achse ist der mLOD-Wert (von -2 bis +3), auf der X-Achse die Multimarkerkarte (von Chromosom 1 bis 22, X und Y) angezeigt. Der höchste Punkt des Graphen markiert die wahrscheinlichste Position des Krankheitslokus. Pfeile ( $\searrow$ ) markieren Kandidatenregionen mit einem mLOD  $> 1,4$  auf den Autosomen.

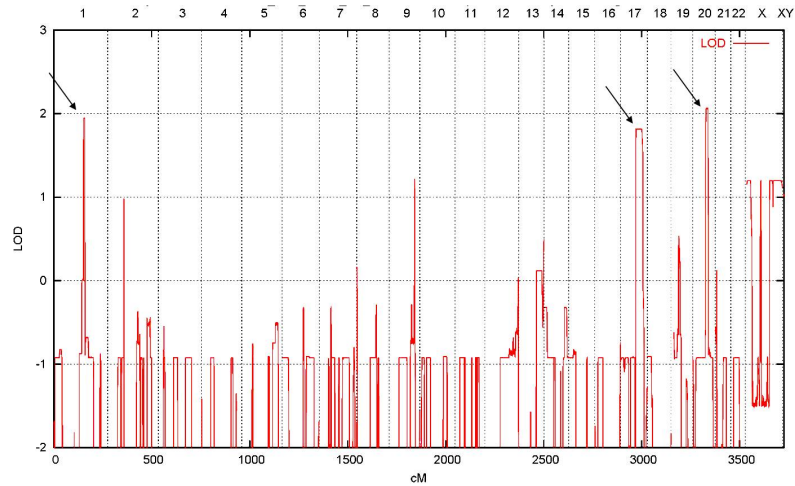


Abbildung 11: *Parametrische Analyse III-1 "unknown"*.

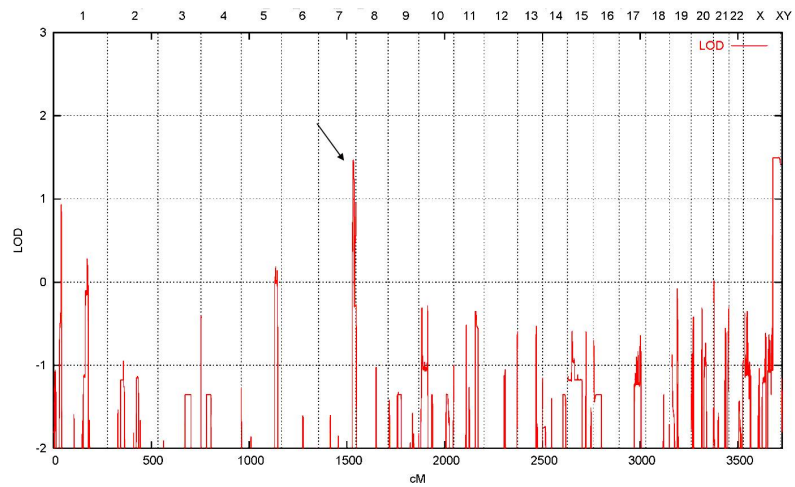
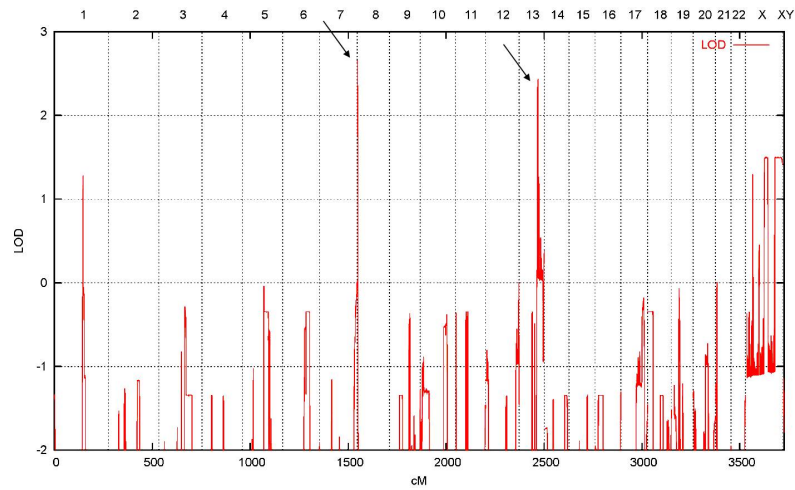
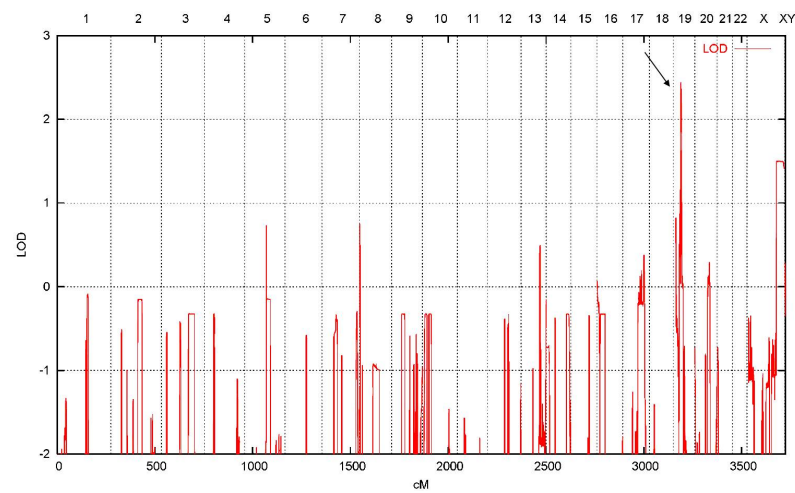
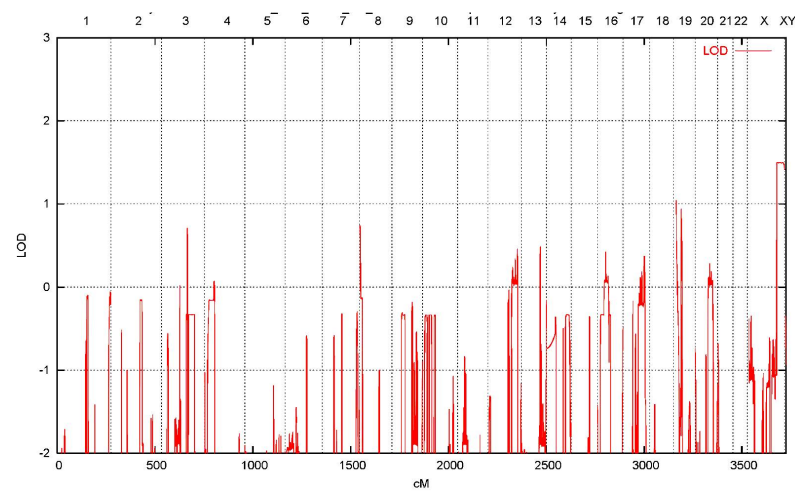


Abbildung 12: *Parametrische Analyse III-5 "unknown"*.

Abbildung 13: *Parametrische Analyse III-6 “unknown”*.Abbildung 14: *Parametrische Analyse III-7 “unknown”*.Abbildung 15: *Parametrische Analyse III-9 “unknown”*.

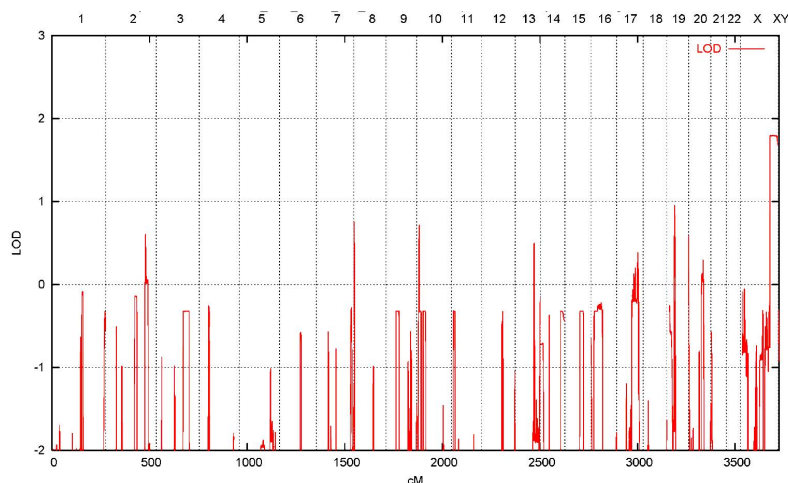


Abbildung 16: *Parametrische Analyse III-10 "unknown"*.

### Person III-1 "unknown"

Bei Berechnung der Analyse mit Person III-1 auf "unknown" gesetzt, entsprachen drei Regionen auf den Autosomen den festgelegten Kriterien (vgl. Abbildung 11). Für 20q wurde der höchste mLOD von 2,066 erreicht, gefolgt von 1q mit einem mLOD von 1,949 und 17q mit einem mLOD von 1,819. Diese drei Regionen zeigten auch in der nicht-parametrischen Analyse die höchsten Werte für  $Z_{\text{mean}}$ : 6,31 für 20q, 5,46 für 1q und 4,75 für 17q. Die unter Annahme 1 identifizierte Kandidatenregion Xq27.2-q28 erreichte in der Analyse mit Person III-1 "unknown" einen mLOD von 1,199.

### Kopplungsregion 1q

Die Haplotypenanalyse zeigte für die Region mit Kopplungssignal auf 1q, dass mit Ausnahme der Person III-1 alle erkrankten Frauen und obligaten Überträgerinnen den krankheitsspezifischen Haplotypen tragen (vgl. Abbildung 17). Person III-1 war in der Analyse auf "unknown" gesetzt worden; sie trägt nicht den Risikohaplotypen und stellt einen sporadischen Brustkrebs in der Familie dar. Die Genotypisierung von drei Mikrosatelliten-Markern in dieser Region bestätigte den krankheitsassoziierten Haplotypen und konnte diesen weiter eingrenzen. Der Tetranukleotidrepeat GATA43A04 liegt in dem Bereich zwischen dem flankierenden und dem folgenden SNP-Marker, die die Kopplungsregion 1q-zentromerwärts begrenzen. Ein Rekombinationsereignis konnte für III-6 und II-7 zwischen GATA43A04 und dem folgenden Marker SNP\_A-1740449 nachvollzogen werden. Das Allel mit 157 Basen für den Dinukleotidrepeat NIB1152 und das Allel mit 177 Basen für AFMA339WH1 ist in allen Trägerinnen des Risikohaplotypen zu finden.

Die Kandidatenregion überspannt einen Bereich von 4,39 Mb in der Region 1q23.1-q23.3. Als flankierender Marker in Richtung Zentromer konnte GATA43A04 identifiziert werden. Telomerwärts wird der Kopplungslokus von SNP\_A-1673316 eingegrenzt. Eine Rekombination konnte zwischen diesem Marker und SNP\_A-1652430 nachvollzogen werden.

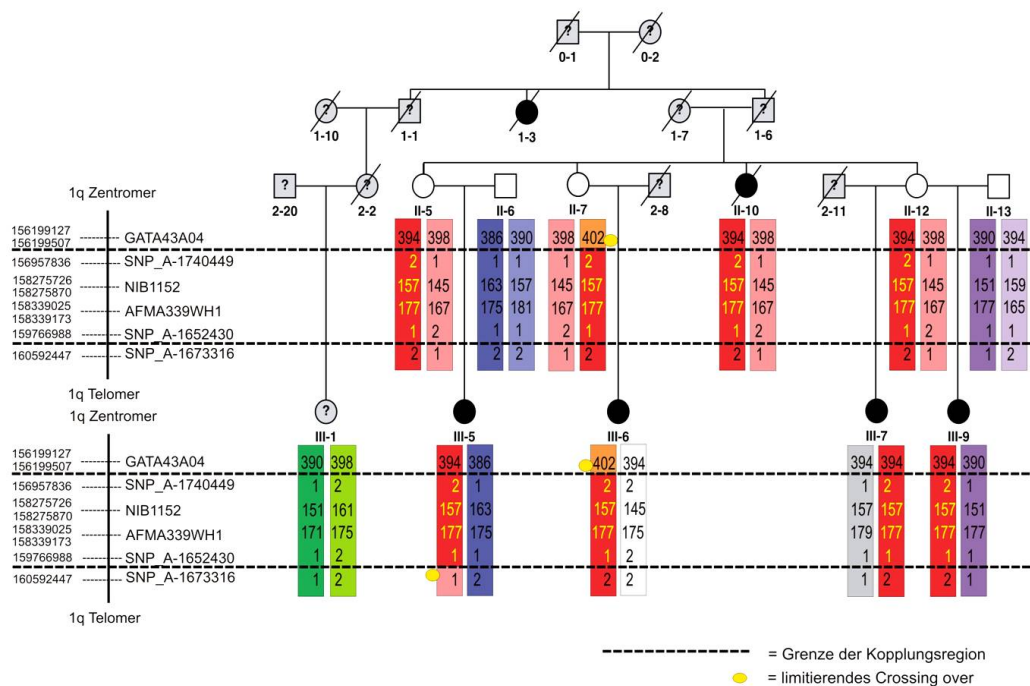


Abbildung 17: **Haplotypenanalyse 1q: Individuum III-1 "unknown"**.

Der Haplotyp in gelber Schrift und rotem Hintergrund segregiert mit der Erkrankung. GATA43A04 und SNP\_A-1673316 sind nicht mehr mit der Erkrankung gekoppelt und stellen die Grenzen der Kandidatenregion 1q dar.

### Kopplungsregion 17q

Die genauere Analyse der Kopplungsregion auf 17q ergab, dass außer Individuum III-1 alle betroffenen Frauen und obligaten Überträgerinnen den rot markierten Risikohaplotypen tragen (vgl. Abbildung 18). Die Person III-1, die in der Analyse auf "unknown" gesetzt worden war, konnte als Phänokopie identifiziert werden. Sie trägt nicht den krankheitsassoziierten Haplotypen. Die Kopplungsregion auf 17q unter Annahme 2 entspricht genau derjenigen unter der Annahme 1 (vgl. Abschnitt 3.1.1). Für den Dinukleotidrepeat AFM338XG5 ist das Allel mit 219 Basen und für den Tetranukleotidrepeat GATA49C09 das Allel mit 469 Basen mit der Erkrankung assoziiert.

Die Kandidatenregion überspannt einen Bereich von 24,08 Mb in der Region 17q21.33-q25.1. SNP\_A-1715757 ist der zentromerwärts, SNP\_A-1656243 der telomerwärts flankierende Marker der Kopplungsregion.



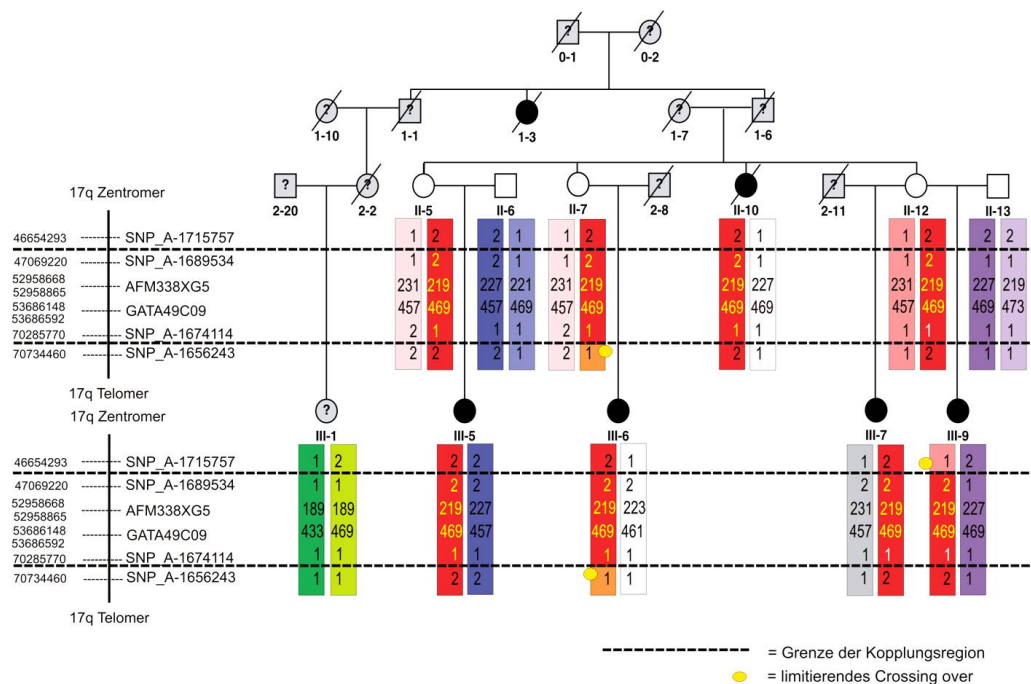


Abbildung 18: **Haplotypenanalyse 17q: Individuum III-1 "unknown".**

Der Haplotyp in gelber Schrift und rotem Hintergrund segregiert mit der Erkrankung, die Allele in weißer Schrift gehören zu Markern, die nicht informativ waren. SNP\_A-1715757 und SNP\_A-1656243 sind nicht mehr mit der Erkrankung gekoppelt und stellen die Grenzen der Kandidatenregion 17q dar.

### Kopplungsregion 20q

Die Haplotypenanalyse der Genotypisierungsdaten zeigte für die potenzielle Kandidatenregion auf 20q, dass mit Ausnahme der Person III-1 alle erkrankten Frauen und obligaten Überträgerinnen den krankheitsassoziierten Haplotypen tragen (vgl. Abbildung 19). Person III-1 war in der Analyse auf "unknown" gesetzt worden; sie trägt nicht den Risikohaplotypen und stellt einen sporadischen Brustkrebs in der Familie dar. Innerhalb der Region des krankheitsspezifischen Haplotypen wurden zwei Mikrosatelliten-Marker einer Fragmentlängenanalyse unterzogen. Für den Dinukleotidrepeat AFMB298YH5 ist das Allel mit 325 Basen, für den Trinukleotidrepeat das Allel mit 242 Basen in der Familie mit der Erkrankung assoziiert.

Die Kopplungsregion überspannt einen Bereich von 8,44 Mb in der Region 20q12-q13.2. SNP\_A-1659628 ist der flankierende Marker der Kandidatenregion in Richtung Zentromer. Ein Rekombinationsereignis konnte zwischen diesem SNP und dem folgenden SNP\_A-1720689 nachvollzogen werden. Telomerwärts wird die Kopplungsregion von SNP\_A-1722314 eingegrenzt. Ein Crossing-over entstand zwischen diesem Marker und SNP\_A-1743231.

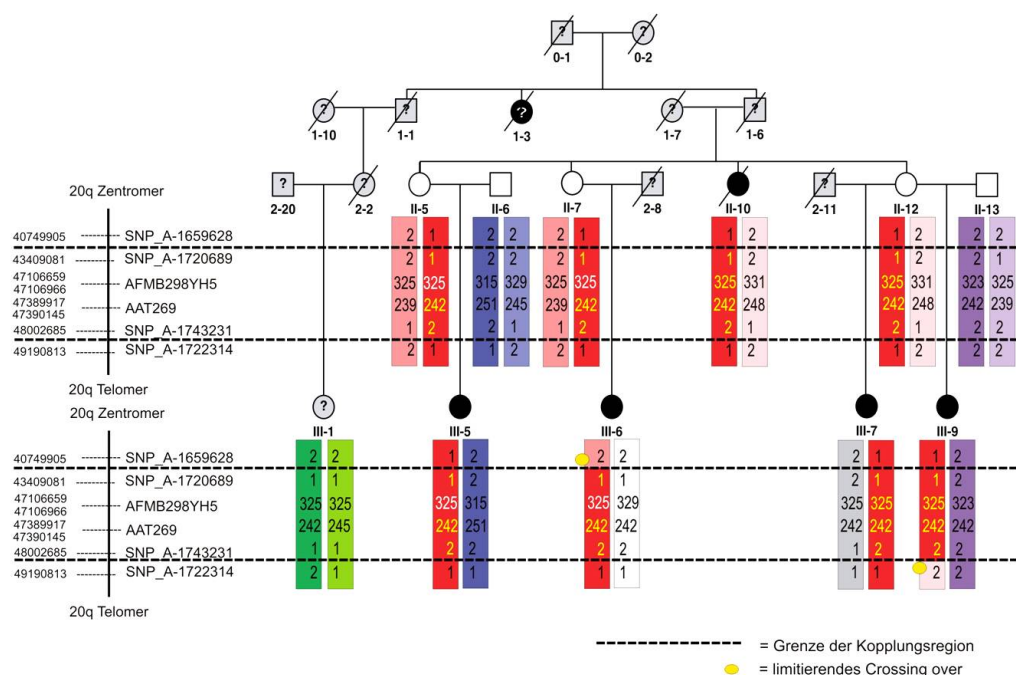


Abbildung 19: **Haplotypenanalyse 20q: Individuum III-1 "unknown".**

Der Haplotyp in gelber Schrift und rotem Hintergrund segregiert mit der Erkrankung, die Allele in weißer Schrift gehören zu Markern, die nicht informativ waren. SNP\_A-1659628 und SNP\_A-1722314 sind nicht mehr mit der Erkrankung gekoppelt und stellen die Grenzen der Kandidatenregion 20q dar.

### Person III-5 "unknown"

In der parametrische Kopplungsanalyse mit Individuum III-5 auf "unknown" gesetzt erreichte nur eine chromosomale Region auf den Autosomen einen mLOD  $> 1,4$  (vgl. Abbildung 12). Ein Locus auf 7q erzielte mit einem mLOD von 1,465 den höchsten Wert. In der nicht-parametrischen Analyse ergab diese Region den maximalen Wert von 6,19 für  $Z_{\text{mean}}$ . Die auf dem X-Chromosom identifizierte Kopplungsregion mit einem mLOD von 1,498 ist identisch der in Abschnitt 3.1.1 beschriebenen Kandidatenregion.

### Kopplungsregion 7q

Die genauere Analyse der potenziellen Kandidatenregion auf 7q ergab, dass mit Ausnahme der Person III-5 alle erkrankten Frauen und obligaten Überträgerinnen den krankheitsassoziierten Haplotypen tragen (vgl. Abbildung 20). Individuum III-5 war in der Analyse auf "unknown" gesetzt worden. Die Haplotypenanalyse ergab, dass III-5 in dieser Region von der Mutter nicht den rot markierten Risikohaplotypen, sondern den rosa markierten Haplotypen geerbt hat. Die Erkrankung Brustkrebs ist für III-5 auf eine sporadische Genese zurückzuführen. Die Genotypisierung von zwei Mikrosatelliten-Markern in dieser Region bestätigte den krankheitsassoziierten Haplotypen und konnte diesen nicht weiter eingrenzen. Das Allel mit 281 Basen für den Dinukleotidrepeat AFMA053XB5 bzw. das Allel mit 257 Basen für den Tetranukleotidrepeat MFD442-GTTT002 findet sich bei den Trägerinnen des Risikohaplotypen.

Die Kandidatenregion überspannt einen Bereich von 2,49 Mb in der Region 7q36.2-q36.3. SNP\_A-1714999 ist der zentromerwärts flankierende Marker der Kopplungsregion. In Richtung Telomer wird der mögliche Krankheitslokus von SNP\_A-1754544 eingegrenzt.

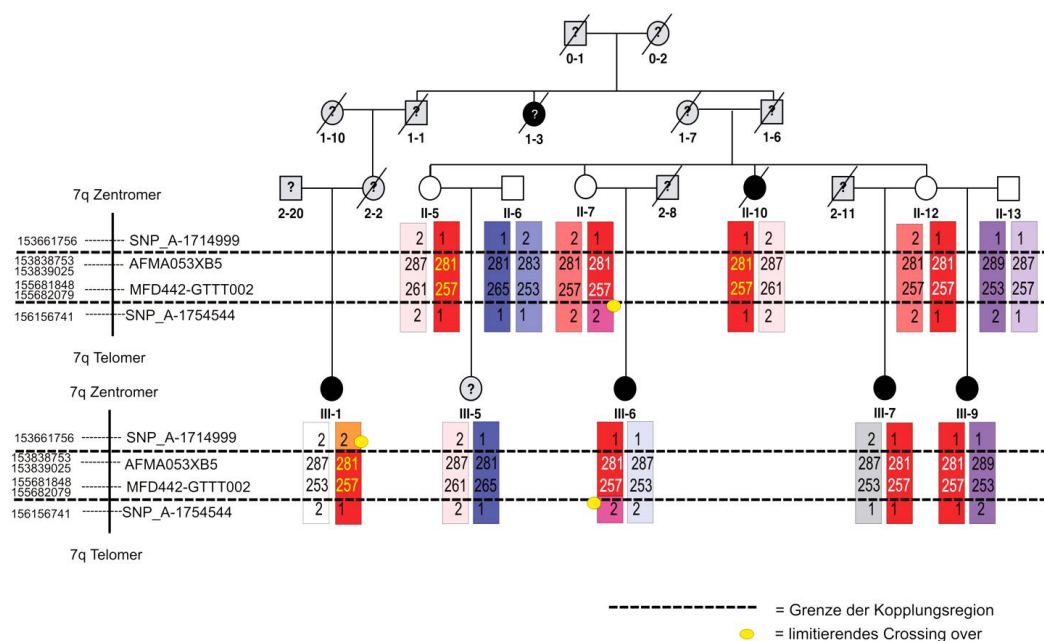


Abbildung 20: **Haplotypenanalyse 7q: Individuum III-5 “unknown”.**

Der Haplotyp in gelber Schrift und rotem Hintergrund segregiert mit der Erkrankung, die Allele in weißer Schrift gehören zu Markern, die nicht informativ waren. SNP\_A-1714999 und SNP\_A-1754544 sind nicht mehr mit der Erkrankung gekoppelt und stellen die Grenzen der Kandidatenregion 7q dar.

### Person III-6 “unknown”

Bei Berechnung der Analyse mit Person III-6 auf “unknown” gesetzt, entsprachen zwei Regionen auf den Autosomen den festgelegten Kriterien (vgl. Abbildung 13). Ein Bereich auf 8p erreichte den höchsten mLOD von 2,654, gefolgt von einem mLOD von 2,429 auf 13q. Diese zwei Regionen zeigten auch in der nicht-parametrischen Analyse die höchsten Werte für  $Z_{\text{mean}}$ : 8,44 für 8p und 5,89 für 13q. Die unter Annahme 1 identifizierte Kandidatenregion auf dem X-Chromosom erreichte in der Analyse mit Individuum III-6 “unknown” einen mLOD von 1,498.

### Kopplungsregion 8p

Die Haplotypenanalyse zeigte für die Kopplungsregion auf 8p, dass mit Ausnahme der Personen III-6 und II-7 alle erkrankten Frauen und obligaten Überträgerinnen den roten Risikohaplotypen tragen (vgl. Abbildung 21). Person III-6, die in der Analyse auf “unknown” gesetzt worden war, konnte als Phänokopie identifiziert werden. Die Haplotypenanalyse ergab, dass bereits die Mutter II-7 keine Trägerin des roten Risikohaplotypen ist. Für den krankheitsassoziierten Haplotypen wurden zwei Mikrosatelliten-Marker einer Fragmentlängenanalyse unterzogen. Beide Marker lagen in dem Bereich, der eventuell weiter eingegrenzt werden konnte. Das Allel mit 162 Basen für den Dinukleotidrepeat  $21 \times \text{TG}$  segregiert in II-5, II-10, II-12, III-1, III-5, III-7 und III-9 mit der Erkrankung. Der Dinukleotidrepeat AFM143XD8 diente der weiteren Eingrenzung der Kandidatenregion auf 8p. Für Person III-9 konnte zwischen diesem Marker und dem Marker  $21 \times \text{TG}$  ein Rekombinationsereignis nachvollzogen werden.

Die Kopplungsregion überspannt einen Bereich von 2,12 Mb in der Region 8p23.3. AFM143XD8 ist der zentromerwärts flankierende Marker der Kopplungsregion.

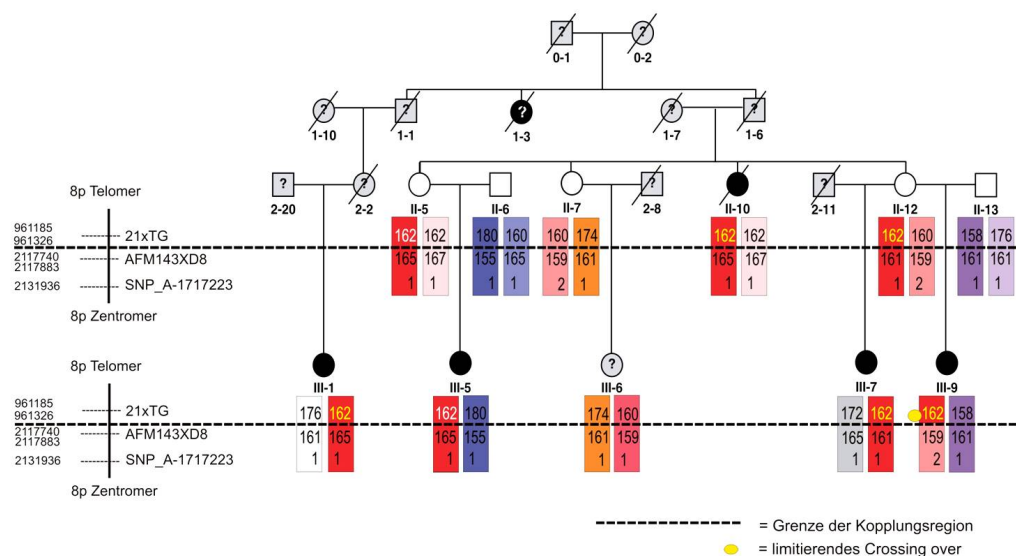


Abbildung 21: **Haplotypenanalyse 8p: Individuum III-6 "unknown"**.

Der Haplotyp in gelber Schrift und rotem Hintergrund segregiert mit der Erkrankung, die Allele in weißer Schrift gehören zu Markern, die nicht informativ waren. AFM142XD8 ist nicht mehr mit der Erkrankung gekoppelt und stellt die Grenze der Kandidatenregion 8p dar.

### Kopplungsregion 13q

Die genauere Analyse der potenziellen Kandidatenregion auf 13q ergab, dass mit Ausnahme der Personen III-1, III-6 und II-7 alle erkrankten Frauen und obligaten Überträgerinnen den krankheitsassoziierten Haplotypen tragen (vgl. Abbildung 22). Person III-6 war in der Analyse auf "unknown" gesetzt worden. Die Haplotypenanalyse zeigte, dass bereits die Mutter II-7 keine Überträgerin des rot markierten Risikohaplotypen ist und Person III-6 somit an einem sporadischen Brustkrebs erkrankte. Individuum III-1 konnte ebenfalls als Phänokopie identifiziert werden. Die Genotypisierung von drei Mikrosatelliten-Markern in der Kopplungsregion 13q bestätigte mit einer Ausnahme den krankheitsspezifischen Haplotypen. Für den Dinukleotidrepeat AFM338WD9 ist das Allel mit 175 Basen, für den Dinukleotidrepeat AFM155YB6 das Allel mit 246 Basen und für den Trinukleotidrepeat das Allel mit 482 Basen mit dem Risikohaplotypen assoziiert. Person II-10 ist homozygot mit 244 Basen für den Marker AFM155YB6 und trägt hier folglich nicht das Risikoallel. Dieses Ergebnis kann einen Genotypisierungsfehler widerspiegeln, durch schwierige Auswertung des Dinukleotidrepeats, der Überlagerungen durch Stotterbanden aufweist, bedingt sein sowie eine Keimbahnmutation in Person 1-6 darstellen.

Die Kandidatenregion überspannt einen Bereich von 17,66 Mb in der Region 13q32.1-q34. Als flankierender Marker der Kandidatenregion konnte SNP\_A-1699107 identifiziert werden. Zwischen diesem Marker und dem folgenden SNP\_A-1674201 konnte ein Rekombinationsereignis in der Familiensituation nachvollzogen werden. Telomerwärts konnte kein Rekombinationsereignis gefunden werden, das die Kopplungsregion begrenzt.

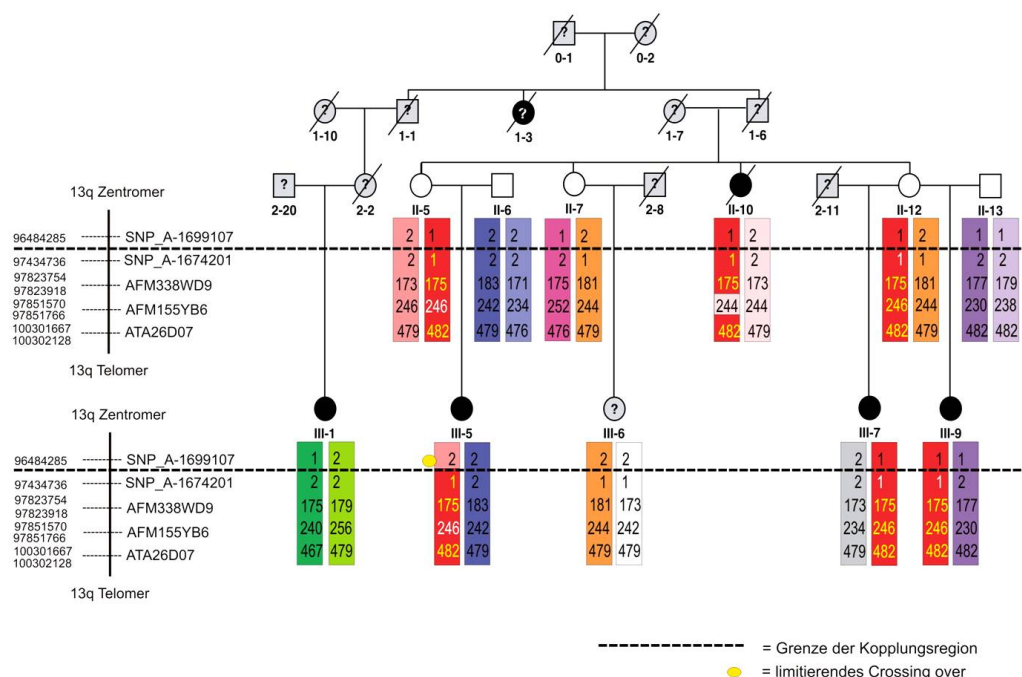


Abbildung 22: **Haplotypenanalyse 13q: Individuum III-6 “unknown”.**

Der Haplotyp in gelber Schrift und rotem Hintergrund segregiert mit der Erkrankung, die Allele in weißer Schrift gehören zu Markern, die nicht informativ waren. SNP\_A-1699107 ist nicht mehr mit der Erkrankung gekoppelt und stellt die Grenze der Kandidatenregion 13q dar.

### Person III-7 “unknown”

In der parametrischen Kopplungsanalyse mit Individuum III-7 auf “unknown” gesetzt erreichte nur eine chromosomale Region auf den Autosomen einen  $mLOD > 1,4$  (vgl. Abbildung 14). Ein Locus auf Chromosom 19 erzielte mit einem  $mLOD$  von 2,444 den höchsten Wert. In der nicht-parametrischen Analyse ergab diese Region den maximalen Wert von 7,03 für  $Z_{mean}$ . Die auf dem X-Chromosom identifizierte Kopplungsregion mit einem  $mLOD$  von 1,499 ist identisch der in Abschnitt 3.1.1 beschriebenen Kandidatenregion.

### Kopplungsregion 19

Die Haplotypenanalyse der Genotypisierungsdaten zeigte für die Kopplungsregion auf Chromosom 19, dass mit Ausnahme der Personen III-1 und III-7 alle erkrankten Frauen und obligaten Überträgerinnen den roten Risikohaplotypen tragen (vgl. Abbildung 23). Individuum III-7 war in der Analyse auf “unknown” gesetzt worden. Die Haplotypenanalyse ergab, dass III-7 in dieser Region von der Mutter nicht den rot markierten Risikohaplotypen, sondern den rosa markierten Haplotypen geerbt hat. Die Personen III-7 und III-1 konnten als Phänokopien identifiziert werden. Für den krankheitsassoziierten Haplotypen wurden drei Mikrosatelliten-Marker einer Fragmentlängenanalyse unterzogen, die diesen bestätigten und zu einer weiteren Eingrenzung der Kandidatenregion führten. Der Tetranukleotidrepeat UT705 liegt in dem Bereich zwischen dem flankierenden und dem folgenden SNP-Marker, die die Kopplungsregion 19p-telomerwärts begrenzen. Ein limitierendes Crossing-over konnte für III-9 zwischen UT705 und dem folgenden Marker SNP\_A-1704341 nachvollzogen werden. Das Allel mit 156 Basen für den Trinukleotidrepeat ATA37G08 und das Allel mit 246 Basen für den Tetranukleotidrepeat GGAA21A04

ist in allen Trägerinnen des Risikohaplotypen zu finden.

Die Kopplungsregion überspannt einen Bereich von 25,38 Mb in der Region 19p13.2-q12. UT705 ist der 19p-telomerwärts, SNP\_A-1666276 der 19q-telomerwärts flankierende Marker der Kandidatenregion.

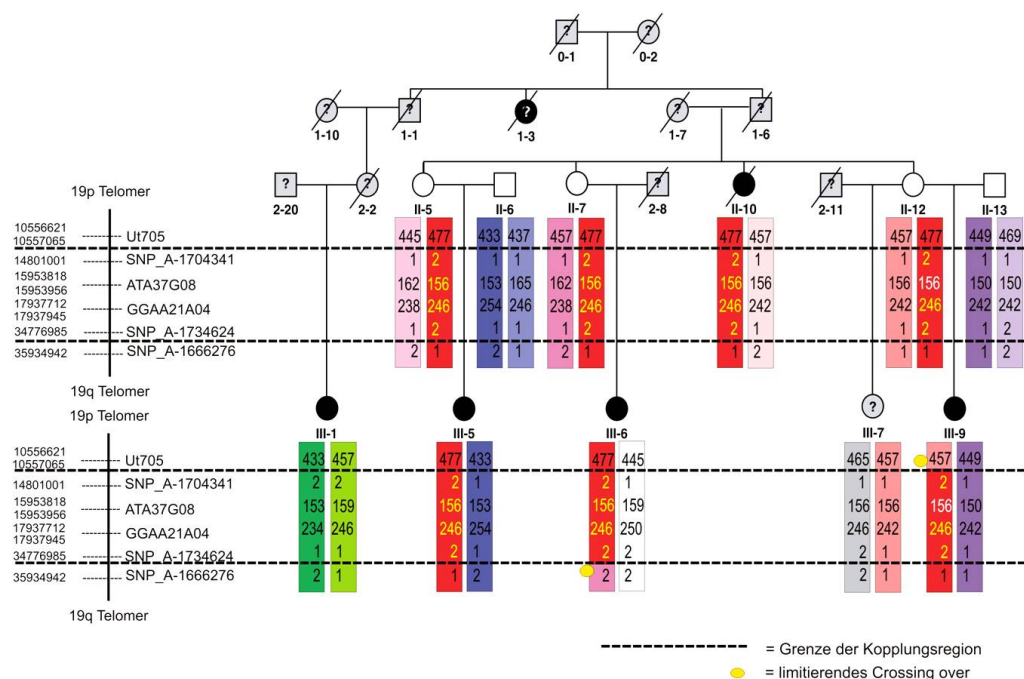


Abbildung 23: **Haplotypenanalyse 19: Individuum III-7 “unknown”.**

Der Haplotyp in gelber Schrift und rotem Hintergrund segregiert mit der Erkrankung, die Allele in weißer Schrift gehören zu Markern, die nicht informativ waren. UT705 und SNP\_A-1666276 sind nicht mehr mit der Erkrankung gekoppelt und stellen die Grenzen der Kandidatenregion auf Chromosom 19 dar.

### Person III-9 bzw. II-10 “unknown”

Bei Berechnung der Analyse mit Individuum III-9 bzw. II-10 auf “unknown” gesetzt erreichte keine Region auf den Autosomen einen  $mLOD > 1,4$  (vgl. Abbildung 15 und 16). Die unter Annahme 1 identifizierte Kandidatenregion auf dem X-Chromosom erreichte bei der Berechnung mit Person III-9 “unknown” einen  $mLOD$  von 1,499 bzw. mit Person II-10 “unknown” einen  $mLOD$  von 1,798.

### 3.1.3 Zusammenstellung der Kandidatenregionen

Kandidaten-Region	Individuum "unknown"	flankierende Marker	physikalische Position (nach hg18)	mLOD	Kopplungs- bereich (in Mb)
<b>Xq27.2-q28</b>	/	SNP_A-1687176 <i>Xq Telomer</i>	140384457 154913754	~ 1,8	14,53
<b>17q21.33-q25.1</b>	/	SNP_A-1715757 SNP_A-1656243	46654293 70734460	~ 1.35	24,08
<b>1q23.1-q23.3</b>	III-1	GATA43A04 SNP_A-1673316	156199127 160592447	<b>1.949</b>	4,39
<b>17q21.33-q25.1</b>	III-1	SNP_A-1715757 SNP_A-1656243	46654293 70734460	<b>1.819</b>	24,08
<b>20q12-q13.2</b>	III-1	SNP_A-1659628 SNP_A-1722314	40749905 49190813	<b>2.066</b>	8,44
<b>7q36.2-q36.3</b>	III-5	SNP_A-1714999 SNP_A-1754544	153661756 156156741	<b>1.465</b>	2,49
<b>8p23.3</b>	III-6	<i>8p Telomer</i> AFM142XD8	1 2117883	<b>2.654</b>	2,12
<b>13q32.1-q34</b>	III-6	SNP_A-1699107 <i>13q Telomer</i>	96484285 114142980	<b>2.429</b>	17,66
<b>19p13.2-q12</b>	III-7	UT705 SNP_A-1666276	10556621 35934942	<b>2.444</b>	25,38

## 3.2 Kandidatengene

In dieser Arbeit wurde mit Hilfe der UCSC Genome Bioinformatics Datenbank in der Karte des humanen Genoms hg18 vom März 2006 nach den Genen gesucht, die in den acht Kandidatenregionen liegen. Es wurden nur Gene mit einem RefSeq-Status berücksichtigt. Für die Kopplungsregion Xq27.2-q28 wurden **123** Gene, für 17q21.33-q25.1 **174** Gene, für 1q23.1-q23.3 **96** Gene, für 20q12-q13.2 **104** Gene, für 7q36.2-q36.3 **9** Gene, für 8p23.3 **11** Gene, für 13q32.1-q34 **71** Gene und für 19p13.2-q12 **278** Gene in die Kandidatengenanalyse miteinbezogen. In den Tabellen 3 – 10 des Anhangs sind die insgesamt 866 Gene in den acht Kandidatenregionen mit Angabe ihres Gennamens, der RefSeq-Nummer, des Transkriptionsstartpunkts und -endpunkts, der Exonzahl, der Genbeschreibung und einer kurzen Funktionsbeschreibung alphabetisch gelistet.

Um die Gene in den Kandidatenregionen nach ihrer Funktion und möglichen Rolle als Kandidatengene für die Erkrankung Brustkrebs zu untersuchen, wurden diese in ein Rangsystem mit den Werten 1, 2, 3 oder 4 eingeordnet. Hierbei wurden die acht Kandidatenregionen unabhängig von ihrem jeweiligen erreichten mLOD-Wert als gleichwertige Regionen für die Lokalisation des genetischen Defekts betrachtet. Rang 1 erhielten nur Gene, die bereits als mögliche Tumorsuppressorgene bzw. Onkogene beschrieben worden sind und für die eine Assoziation mit Brustkrebs (Amplifikation, Expressionsveränderung, LOH in Brustkrebs) nachgewiesen worden ist. Der Rang 2 zeigt an, dass für dieses Gen entweder eine Assoziation mit Brustkrebs beschrieben worden ist oder es eine mögliche Rolle in der Karzinogenese spielen könnte. Das sind zum einen Gene, die als Tumorsuppressoren oder Onkogene wirken könnten und zum anderen Gene, die als DNA-Reparaturgene, Apoptosegene oder Telomerasegene zu Tumorgenen transformieren könnten. Gene mit Rang 3 wurden aufgrund ihrer bisher bekannten Funktion als unwahrscheinliche Kandidatengene für Brustkrebs ausgeschlossen. Den Rang 4 erhielten Gene, über die aufgrund des fehlenden oder unzureichenden Wissens bezüglich ihrer Funktion

keine Aussage gemacht werden konnte. Diese Gene können jedoch als Kandidatengene für die Brustkrebs susceptibility in der vorliegenden Familie nicht ausgeschlossen werden.

Durch diese auf Datenbank- und Literaturrecherche basierende Kandidatengenanalyse konnte 56 Genen in den acht Kopplungsregionen ein Rang 1 oder 2 zugewiesen werden. Zusätzlich wurden in den Kandidatenregionen diejenigen Gene identifiziert, die nach der Netzwerkmodell-Strategie von Pujana et al. (2007) [45] als interessante Gene für Brustkrebsprädisposition beschrieben worden sind. Dies war für neun Gene zutreffend (vgl. Abschnitt 4). Die Gesamtzahl von 866 Genen in den acht Kopplungsregionen konnte durch diese systematische Funktionsanalyse auf 62 mögliche Kandidatengene eingeschränkt werden. Für die Kandidatenregion Xq27.2-q28 entsprachen **2** Gene, für 17q21.33-q25.1 **18** Gene, für 1q23.1-q23.3 **10** Gene, für 20q12-q13.2 **11** Gene, für 7q36.2-q36.3 **2** Gene, für 8p23.3 **kein** Gen, für 13q32.1-q34 **7** Gene und für 19p13.2-q12 **12** Gene den Voraussetzungen. Um eine Rangordnung innerhalb der 62 Kandidatengenen für Brustkrebs susceptibility zu erstellen, wurden diese im nächsten Schritt auf das Übereinstimmen mit vier Kriterien überprüft:

- Funktion als mögliches Tumorsuppressorgen bzw. Onkogen
- Assoziation des Gens mit Brustkrebs (Expressionsveränderung, Amplifikation, LOH in Brustkrebs)
- Rolle des Gens in der DNA-Reparatur als kritischen Signalweg für Brustkrebs susceptibility
- Funktion des Gens im Apoptose-Signalweg, in der Proliferation oder in der Zellzykluskontrolle. Positive Wirkung des Gens auf die Proliferation bzw. negative Wirkung auf die Apoptose bei fehlender bzw. herunterregulierter Expression oder überaktivierter Transkription.

In der folgenden Tabelle sind die 62 Kandidatengene gelistet und den vier Kriterien zugeordnet.

	Gen Symbol	Rang	Gen nach Pujana et al. [45]	möglicher/s Tumor-suppressor/ Onkogen	assoziiert mit Brustkrebs	Rolle in DNA-Reparatur	Rolle in Apoptose, Proliferation oder Zellzykluskontrolle
Xq	BRCC3	2				+	+
Xq	TREX2	2				+	+
17q	Sep 04	2					+
17q	APPBP2	2			+		
17q	AXIN2	2			+		+
17q	BCAS3	2			+		
17q	BRIP1	1		+	+	+	
17q	C17orf28	2			+		
17q	ERN1	2					+
17q	FDXR	2					+
17q	HELZ	2			(+)		
17q	LYK5	2		(+)			+
17q	PPM1D	2			+		+
17q	PTRH2	2					+
17q	RAD51C	1	+	+	+	+	



17q	RPS6KB1	2			+		+
17q	SLC9A3R1	1		+	+		+
17q	SUMO2	2				(+)	(+)
17q	TBX2	2			+		
17q	TRIM25	2		(+)	+		
1q	AIM2	2			+		+
1q	CD1C	3	+				
1q	DEDD	2					+
1q	IFI16	2			+		+
1q	NIT1	2		+			+
1q	PEA15	2			+		+
1q	PEX19	2					+
1q	PYHIN1	1		+	+		+
1q	SDHC	2		+			+
1q	SUMO1	2				(+)	(+)
20q	ADNP	2			+		+
20q	BCAS4	2			+		
20q	C20orf10	2					+
20q	CSE1L	2					+
20q	MYBL2	4	+				+
20q	NCOA3	2			+		
20q	STK4	2		+			+
20q	TP53RK	2					+
20q	UBE2C	3	+				
20q	UBE2V1	2				+	+
20q	WISP2	2			+		+
7q	EN2	2		+	+		
7q	PAXIP1	2	+			(+)	+
13q	CUL4A	2		+	+	(+)	
13q	ERCC5	2				+	
13q	GAS6	2			+		+
13q	ING1	1		+	+		+
13q	LAMP1	2			+		
13q	LIG4	2				+	
13q	TFDP1	2	+		+		+
19	CCNE1	2					+
19	DDX39	3	+				
19	GADD45	2		+			+
19	ILF3	4	+				
19	JUNB	2		+			
19	PKN1	2					+
19	PLEKHF1	2					+
19	PRDX2	2					+
19	RAD23A	2				+	
19	SMARCA4	2		+			+
19	UBA52	2				(+)	
19	ZNF43	4	+				

(+) steht für einen indirekten Zusammenhang mit dem entsprechenden Kriterium.

Da es den Rahmen dieser Arbeit überschritten hätte, alle 62 Kandidatengene durch Sequenzanalyse zu untersuchen, wurden durch die Kandidatengenanalyse systematisch diejenigen Gene identifiziert, die nach den angesetzten Kriterien am wahrscheinlichsten für den Brustkrebs in der untersuchten Familie verantwortlich sein könnten. Die fünf

Gene BRIP1 (17q), RAD51C (17q), SLC9A3R1 (17q), PYHIN1 (1q) und ING1 (13q) erfüllten mindestens drei der angesetzten Kriterien und stellten somit die interessantesten Kandidatengene dar. Die übrigen 57 Gene erfüllten weniger als drei Kriterien und wurden, als zweitrangig eingestuft, nicht weiter analysiert.

### 3.3 Sequenzierung

Im Rahmen dieser Arbeit wurden die vier möglichen Kandidatengene RAD51C, PYHIN1, ING1 und SLC9A3R1 durch eine Sequenzanalyse näher untersucht. Das Gen BRIP1 ist bereits im Vorfeld sequenziert worden, ohne dass pathogene Mutationen gefunden werden konnte. Die erhaltenen Sequenzen wurden mit der Zielsequenz des hg18 vom März 2006 aus der öffentlichen UCSC Genome Bioinformatics Datenbank verglichen und auf Abweichungen hin untersucht.

Da die Analyse des Stammbaums einen autosomal dominant vererbten genetischen Defekt vermuten lässt, kann man davon ausgehen, dass alle Betroffenen und Überträger des Gendefekts die verursachende Mutation nur in einer Kopie ihrer homologen DNA-Sequenzen aufweisen. Gesunde Kontrollpersonen sollten diese Mutation nicht tragen. Unter allen gefundenen Sequenzveränderungen tritt diese Situation jedoch nicht auf.

Das Gen RAD51C besteht aus 9 kodierenden Exonen, die für III-5, II-5 und II-6 sequenziert wurden. In Exon 1 wurde eine Abweichung zur Vergleichssequenz hg18 gefunden:

SNP	II-5	II-6	III-5	hg18	Ancestrales Allel	Allele	Average Het. +/- std. error
rs12946397	T/T	C/C	C/T	C/C	C	C,T	0.323 +/-0.239

Dabei handelt es sich jedoch um einen polymorphen SNP. Keine weiteren Sequenzveränderungen wurden gefunden.

Das Gen ING1 enthält 2 kodierende Exone, die für III-9, II-12 und II-13 sequenziert wurden. In den Isoformen b und c des Exon 1 konnten insgesamt drei Sequenzveränderungen im Vergleich mit der hg18-Sequenz gefunden werden:

SNP	II-12	II-13	III-9	hg18	Ancestrales Allel	Allele	Average Het. +/- std. error
rs12875067	A/G	G/G	G/G	A/A	G	G,A	0.447 +/-0.154
rs7338333	G/G	G/G	G/G	T/T	G	G,T	unbekannt
rs9555726	T/T	T/T	T/T	C/C	C	C,T	0.491 +/-0.065

Dabei handelt es sich jedoch jeweils um einen polymorphen SNP. Keine weiteren Sequenzveränderungen wurden gefunden.

Das Gen PYHIN1 besteht aus 9 kodierenden Exonen, die für III-5, II-5 und II-6 und Exon 8 zusätzlich für III-9, II-12 und II-13 analysiert wurden. Es wurde keine Sequenzveränderung gefunden.

Das Gen SLC9A3R1 besteht aus 6 kodierenden Exonen, die für III-9, II-12 und II-13 analysiert wurden. Es wurde keine Sequenzveränderung gefunden.

## 4 DISKUSSION

In der vorliegenden Arbeit wurde mittels des positionellen Kandidatengenverfahrens bei einer BRCA1/2-negativen Familie, in der das Mammakarzinom in einer autosomal dominanten Weise segregiert, nach einem neuen "Brustkrebsgen" gesucht.

In der detaillierten Auswertung der Kopplungsdaten, auch unter der Annahme von sporadischen Phänokopien, wurden acht Kandidatenregionen auf acht unterschiedlichen Chromosomen identifiziert. Diese konnten durch die Genotypisierung von Mikrosatelliten-Markern bestätigt werden und die exakten Grenzen der Kopplungsregionen wurden bestimmt. Mittels einer funktionsorientierten Analyse wurde unter den 866 in den Kopplungsregionen liegenden Genen nach möglichen Kandidatengen für die Erkrankung Brustkrebs gesucht. Aus den 62 identifizierten Kandidatengen wurden die fünf geeignetsten (BRIP1, RAD51C, PYHIN1, ING1 und SLC9A3R1) mittels Sequenzanalyse untersucht. Es konnte jedoch keine krankheitsassoziierte Mutation entdeckt werden.

Trotz der systematischen Vorgehensweise konnte im Rahmen dieser Arbeit der genetische Defekt, der in der untersuchten Familie für den familiär auftretenden Brustkrebs ursächlich ist, nicht gefunden werden. Ausgehend von den vorliegenden Ergebnissen zeigen sich jedoch einige Punkte, die einer tieferen Diskussion bedürfen.

Zunächst stellt sich die Frage, ob das positionelle Kandidatengenverfahren die richtige Strategie ist, um das verantwortliche Krankheitsgen zu identifizieren. Ferner muss die Wahl des Vererbungsmodus "autosomal dominant" und die Problematik des vorhandenen Kopplungssignals auf dem X-Chromosom beleuchtet werden. Es besteht zudem die Möglichkeit, dass die in der Kopplungsanalyse identifizierten Kandidatenregionen falsch positive Signale widerspiegeln und somit der verantwortliche Krankheitsloкус nicht herausgefiltert wurde. Hierfür muss auf die Signifikanz der Kopplungsregionen und die statistische Aussagekraft der Familie eingegangen werden. Des Weiteren stellt sich bezüglich der Identifizierung aller Gene in den acht Kandidatenregionen die Frage nach der Vollständigkeit der gelisteten Gene. Obwohl die anschließend durchgeführte Kandidatengenanalyse zu Erfolg versprechenden Ergebnissen führen sollte, ist eine Identifikation und Weiterverfolgung "nicht relevanter" Kandidatengene nicht auszuschließen. Ebenfalls besteht die Möglichkeit, dass das richtige Krankheitsgen ausgeschlossen wurde. Der Methodik der Sequenzierung sind bei der Detektion der krankheitsverursachenden Mutation Grenzen gesetzt. Somit dürfen die fünf sequenzierten Kandidatengene nicht als endgültig ausgeschlossen betrachtet werden, für den in der Familie auftretenden Brustkrebs verantwortlich zu sein. Es ist zudem interessant einen Vergleich zu den in anderen Kopplungsstudien gefundenen Kandidatenregionen anzustellen. Abschließend soll unter Bezugnahme der Kopplungsdaten erörtert werden, ob die zwei großen "Brustkrebsgene" BRCA1 und BRCA2 für die untersuchte Familie wirklich auszuschließen sind.

Eine genetische Disposition wird bei 5-10% aller Brustkrebspatientinnen als Ursache der Erkrankung angenommen. Zurzeit sind einige, aber längst nicht alle beteiligten Genveränderungen und ihre Auswirkungen bekannt. Das gehäufte Vorkommen von Familien mit einer hohen Konzentration an Brustkrebserkrankungen, die negativ für eine Mutation in den Genen BRCA1 und BRCA2 sind, lässt vermuten, dass noch ein oder mehrere autosomal dominant vererbte Brustkrebs-suszeptibilitätsgene mit hoher Penetranz existieren [14]. Die hier vorliegende Arbeit basiert auf der Untersuchung einer einzigen Familie, in der sieben Frauen, mindestens sechs von ihnen vor dem fünfzigsten Lebensjahr, die

Diagnose Brustkrebs erhielten. Anhand dieser Familie mit vielen Erkrankungen in jungen Jahren und dem weitgehenden Ausschluss einer Beteiligung von BRCA1 und BRCA2, lässt sich die Vermutung anstellen, dass ein BRCA1/2-ähnliches Gen für das erhöhte Brustkrebsrisiko verantwortlich ist. Das Auftreten der Erkrankung im Stammbaum lässt einen autosomal dominanten und monogen vererbten genetischen Defekt mit unvollständiger Penetranz vermuten, der das familiäre Mammakarzinom bedingt.

Eine sehr erfolgreiche Strategie zur Identifizierung eines Krankheitsgens mit hoher Penetranz ist die Kopplungsanalyse mit einem sich anschließenden positionellen Klonierungsansatz. Unter diesem Ansatz wurde das positionelle Kandidatengenverfahren gewählt, um bei der vorliegenden Brustkrebs-Familie das verantwortliche Krankheitsgen zu finden. Es ist jedoch zu beachten, dass die familiäre Häufung der Erkrankung auch durch eine Koinkidenz mehrerer Suszeptibilitätsallele mit niedriger Penetranz bedingt sein könnte, die additiv oder multiplikativ das Risiko beeinflussen, und zusammen mit Umwelttrisikofaktoren für Brustkrebs verantwortlich sind [12]. Suszeptibilitätsgene, die mit einem niedrig bis mäßig erhöhten Risiko zu erkranken verbunden sind, können mit der gewählten Methode nicht identifiziert werden, da die Beziehung zwischen Genotyp und Phänotyp sehr viel schwächer ist [14]. Dies zeigt die Problematik der positionellen Kandidatengenanalyse bei Krankheiten mit komplexer Ätiologie, die sich in der Analyse auf einzelne Familien beschränken müssen.

Die Aussagekraft einer Kopplungsanalyse wird von mehreren Faktoren beeinflusst, welche zu einer Fehlinterpretation der in dieser Arbeit identifizierten Kopplungssignale führen können. Das dominante Vererbungsmodell *rare dominant*, das der parametrischen Kopplungsanalyse zugrundegelegt wurde, könnte sowohl in der Festlegung der Parameter (Frequenz und Penetranz) als auch des Vererbungsmodus nicht der Realität entsprechen und zur Identifizierung "falscher" bzw. artifizieller Kopplungssignale geführt haben. Daher war es von Vorteil, sowohl parametrische als auch die vom Vererbungsmodus unabhängige nicht-parametrische Kopplungsanalysen durchzuführen, um falsch positive bzw. falsch negative Ergebnisse zu vermeiden. Die in der parametrischen Kopplungsanalyse identifizierten acht Kandidatenregionen zeigten jeweils auch in der nicht-parametrischen Analyse übereinstimmende Ergebnisse.

In der parametrischen Kopplungsanalyse wurde unter der Annahme ("Annahme 1"), dass allen Erkrankungen derselbe genetische Defekt zugrunde liegt, der höchste mLOD von 1,8 auf dem X-Chromosom erzielt. Das höchste Kopplungssignal auf den Autosomen lag auf 17q ( $mLOD = 1,35$ ). Das hohe Kopplungssignal auf dem X-Chromosom wies auf eine X-gekoppelte Vererbung des Krankheitsgens hin und widersprach somit dem aus der Stammbaumanalyse hervorgehenden autosomal dominanten Vererbungsmodus. X-chromosomale Merkmale sind meist rezessiv mit voller Penetranz in den hemizygoten Männern. Da bei zwei der im Stammbaum aufgeführten Männern, über welche der Gendefekt weitergegeben hätte werden müssen, keine Erkrankung Brustkrebs bekannt ist, erschien ein X-chromosomal Erbgang in dieser Familie nicht wahrscheinlich. Trotz dieser Unstimmigkeiten wurde die Kopplungsregion auf dem X-Chromosom in die weitere Analyse miteinbezogen.

Weiter stellt das Phänomen der inkompletten Penetranz ein Problem dar, da von dem Trägerstatus des Krankheitsallels nicht definitiv auf den Krankheitsstatus geschlossen werden kann. In der Stammbaumanalyse wurden obligate Überträgerinnen des Gendefekts identifiziert, für die noch keine Diagnose Brustkrebs gestellt worden war. Der genetische Defekt wirkt in der untersuchten Familie somit mit unvollständiger Penetranz.

Um dieser Tatsache gerecht zu werden, wurde die Penetranz im Vererbungsmodell der Kopplungsanalyse auf 0,90 festgelegt.

Die Auswertung der Kopplungsdaten wurde zusätzlich unter einer weiteren Annahmen ("Annahme 2") durchgeführt, welche die Problematik eventuell vorkommender Phänokopien adressieren sollte. Diese berücksichtigt, dass eine betroffene Person nicht aufgrund des in der Familie mit der Erkrankung segregierenden genetischen Defekts, sondern aufgrund sporadischer Genese an Brustkrebs erkrankte. Diese Einschränkung in der Auswertung ist in Hinblick auf die hohe Prävalenz von Brustkrebs in der Bevölkerung zu beachten. Die Berechnung erfolgte, indem nacheinander verschiedene Betroffene auf "unknown" gesetzt wurden. Unter dieser Annahme wurden bei der Analyse als mögliche Kandidatenregionen solche chromosomalen Regionen in Betracht gezogen, die in der parametrischen Analyse einen mLOD-Wert  $> 1,4$  auf den Autosomen erzielten. So konnten sieben Kopplungsregionen identifiziert werden, wobei eine Kopplungsregion (17q) unter Annahme 1 und Annahme 2 einander entsprach. Von einer weiteren Annahme, dass zwei oder sogar drei der an Brustkrebs erkrankten Personen aufgrund sporadischer Genese erkrankten und somit Phänokopien darstellen, wurde im Rahmen dieser Arbeit abgesehen. In der Kopplungsanalyse würden nicht angegebene Phänokopien den mLOD-Wert zusätzlich verringern.

Das Signifikanzniveau ist eine wichtige Größe zur Bestimmung des Vorliegens einer signifikanten Kopplung. Während für X-chromosomale Loki ein Signifikanzlevel von 0,01 als ausreichend betrachtet wird, wird bei autosomalen Loki ein Level von 0,001 angesetzt; d.h. das Ergebnis wird als signifikant erachtet, wenn der maximale LOD einen Wert von mindestens 3 (autosomale Loki) bzw. mindestens 2 (X-chromosomale Loki) erreicht [3]. Die acht identifizierten Kopplungsregionen dieser Studie können aufgrund der festgestellten mLOD-Werte von maximal 2,654 (auf 8p) nicht als "klassisch" signifikant gewertet werden.

Die Wahrscheinlichkeit, dass der gesuchte genetische Defekt in der postulierten Kopplungsregion liegt, wird durch den Faktor  $10^X$ , wobei X den mLOD-Wert angibt, ausgedrückt.

Diese Wahrscheinlichkeit ist in der folgenden Tabelle für jede identifizierte Kandidatenregion aufgeführt.

Kopplungsregion	mLOD	Wahrscheinlichkeit $10^X$
<b>"Annahme 1"</b>		
Xq	1,8	63,1
17q	1,35	22,4
<b>"Annahme 2"</b>		
1q	1,949	88,9
17q	1,819	65,9
20q	2,066	116,4
7q	1,465	29,2
8p	2,654	450,8
13q	2,429	268,5
19	2,444	278,0

Die telomerische Region auf Chromosom 8p stellt mit einem mLOD von 2,645 den wahrscheinlichsten Lokus für Kopplung mit dem hypothetischen Krankheitslokus dar. Dieses Kopplungssignal könnte jedoch falsch positiv sein, da telomerische Regionen dazu neigen erhöhte LOD-Werte aufzuweisen [46]. In der Kandidatengenanalyse zeigte sich für diese Kopplungsregion auch kein mögliches Kandidatengen.

Für die statistische Aussagekraft einer Kopplungsanalyse spielt neben der Anzahl der untersuchten Familienmitglieder auch die Struktur der Familie eine Rolle. Ideale Strukturen für eine Kopplungsanalyse sind bei Familien zu finden, in denen drei Generationen mit jeweils acht Kindern vorhanden sind und alle vier Großelternpaare zur Verfügung stehen [3]. Eine Familie mit diesen Voraussetzungen ist allerdings selten und lag auch in dieser Untersuchung nicht vor. Die statistische Aussagekraft der Familie ist demnach gering. Es stehen zu wenige Meiosen zur Verfügung, die erforderlich wären, um einen ausreichend hohen und damit signifikanten LOD-Wert von 3 zu erzielen. Jede Meiose in einer Familie, d.h. jede Möglichkeit zur Rekombination während der Entstehung einer Keimzelle, trägt  $\log(2) \approx 0,3$  zur Höhe des LOD-Werts bei. Aus statistischen Gründen muss vor der Berechnung des erwarteten LOD-Werts eine Meiose pro Familie von der Gesamtzahl subtrahiert werden [3]. Nach dieser Faustregel kann die Kopplungsanalyse der dieser Arbeit zugrundeliegenden relativ kleinen Familie mit 9 Meiosen (II-5, II-7, II-10, II-12, III-1, III-5, III-6, III-7 und III-9) einen maximalen LOD-Wert von etwa  $8 \times 0,3 = 2,4$  erwarten. Zusätzlich verringert das Auftreten von Phänokopien den mLOD. Das Vorliegen nicht-signifikanter Kandidatenregionen ist somit auf die zu geringe Kraft eine signifikante Kopplung zu detektieren zurückzuführen. Um eventuell falsch positive Kopplungsregionen zu identifizieren und damit die Anzahl der Kandidatenregionen einzuschränken, ist es notwendig die Kraft der Analyse zu erhöhen. Dies könnte durch die Einbeziehung der fünften Generation in die genomweite Analyse und die dadurch zur Berechnung des LOD-Werts zur Verfügung stehenden weiteren Meiosen erreicht werden. Die Schwierigkeit für die Berücksichtigung der fünften Generation besteht darin, dass der Phänotyp der Erkrankung Brustkrebs bei diesen Familienmitgliedern aufgrund zu niedrigen Alters noch nicht aufgetreten ist.

Um die Kandidatenregionen und deren genaue Grenzen zu identifizieren, bediente man sich der Haplotypenanalyse. Bei der Haplotypenanalyse wurden die wahrscheinlichsten Haplotypen für die genotypisierten Individuen ermittelt, indem die Allele der Elterngeneration derart den jeweiligen Chromosomen zugeordnet wurden, dass mit einer minimalen Anzahl an Rekombinationen alle Haplotypen der nachfolgenden Generation rekonstruiert werden konnten.

Ist nun ein Marker nicht informativ, können die Allele nicht eindeutig zugeordnet werden; d.h. ist eine Person der Elterngeneration homozygot für einen Marker oder sogar für mehrere aufeinanderfolgende Marker, so kann die Lokalisation eines Rekombinationsergebnisses in dieser Region nicht eindeutig bestimmt werden.

Um die krankheitsassoziierten Haplotypen zu bestätigen und die identifizierten Kandidatenregionen der Kopplungsanalyse zu überprüfen, wurden für jede identifizierte Kopplungsregion Mikrosatelliten-Marker genotypisiert. Die Ergebnisse der Haplotypenanalyse konnten durch den Einsatz der Mikrosatelliten-Marker repliziert und somit bestätigt werden.

In der positionellen Kandidatengenanalyse wurde mit Hilfe der UCSC Genome Bioin-

formatics Datenbank in der Karte des humanen Genoms hg18 vom März 2006 nach den Genen gesucht, die in den acht Kandidatenregionen liegen. Mit diesem Kandidatengenverfahren konnten nur solche Gene analysiert werden, die nach dem heutigen Stand der Forschung in dieser öffentlichen Gendatenbank annotiert sind. Es wurden ferner nur Gene mit einem sogenannten RefSeq-Status berücksichtigt. Dieser gibt Auskunft über die Richtigkeit und Verlässlichkeit der angegebenen Sequenzen und spiegelt das aktuelle Wissen über das Gen und seine Transkripte wieder. Die Status-Level werden *Inferred*, *Model*, *Predicted*, *Provisional*, *Validated* oder *Reviewed* genannt. Insgesamt wurden 866 Gene in den acht Kandidatenregionen in die weitere Analyse einbezogen.

In den Kandidatenregionen lagen sowohl einige hypothetische Proteine, als auch zahlreiche bisher unbekannte ESTs (*expressed sequenced tags*). Diese ESTs wurden zwar als kodierende DNA-Stücke identifiziert, ihnen konnte jedoch bisher noch kein Gen zugeordnet werden. Da diese meist keinen RefSeq-Status besitzen, wurden sie in der Kandidatengenanalyse nicht berücksichtigt. Hypothetische Proteine sind, sofern sie eine RefSeq-Statuszuordnung besitzen, in den Tabellen 3-10 des Anhangs aufgeführt. Da über ihre Funktion jedoch meist sehr wenig bekannt ist, konnte keine Aussage über eine mögliche Rolle als Brustkrebs-verursachendes Gen gemacht werden. Das gleiche gilt für einige weitere Gene, deren Funktion noch unzureichend geklärt ist. Aufgrund der Unvollständigkeit bekannter Gensequenzen und der unbekannten Funktion vieler Gene besteht die Möglichkeit, dass das krankheitsverursachende Gen bereits bei der Auswahl der Gene in den Kandidatenregionen nicht berücksichtigt wurde.

Unter der Annahme, dass der tatsächliche Krankheitslokus in der durchgeführten Kopplungsanalyse zu detektieren ist, wurde eine Strategie zur Identifizierung von möglichen Kandidatengen entwickelt. Diese Reduktion der Genanzahl war erforderlich, da es im Rahmen dieser Arbeit nicht möglich war die 866 Gene in den acht Kandidatenregion auf eine pathogene Mutation zu sequenzieren. Kandidatengene stellen solche Gene dar, die in einer Kopplungsregion lagen und deren Funktion oder vielmehr deren Defekt mit dem Phänotyp der Erkrankung Brustkrebs in Zusammenhang gebracht werden konnte.

Die in dieser Arbeit durchgeführte Kandidatengenanalyse basierte auf einer Datenbank- und Literaturrecherche, in der alle 866 Gene nach einer möglichen Rolle als Kandidatengene für Brustkrebs bewertet wurden. Hierfür wurde jedem Gen ein Rang von eins bis vier zugeordnet (vgl. Abschnitt 3.2). Der Rang 1 zeigte an, dass dieses Gen ein potenzielles Kandidatengen für Brustkrebs darstellt. Rang 2 erhielten Gene, die eine mögliche Rolle in der Krebsentstehung im Sinne eines Tumorgens spielen könnten oder für die eine Assoziation mit Brustkrebs beschrieben worden war. Gene mit Rang 3 wurden als unwahrscheinliche Kandidatengene für Brustkrebs betrachtet, während für Gene mit Rang 4 aufgrund des fehlenden oder unzureichenden Wissens über deren Funktion keine Aussage gemacht werden konnte. 396 Gene fielen in die Kategorie mit Rang 4 und konnten somit als Kandidatengene für die Brustkrebsprädisposition in der vorliegenden Familie nicht ausgeschlossen werden. Einen Rang 1 oder 2 erhielten nach dieser Recherche 56 Gene, die als interessante Gene für Brustkrebs-suszeptibilität in die Liste der Kandidatengen aufgenommen wurden.

Ferner wurden alle 866 Gene nach der Studie von Pujana et al. (2007) [45] bewertet. Diese arbeitete mit einer Netzwerkmodell-Strategie zur Identifizierung von Genen, die möglicherweise mit einem erhöhten Brustkrebsrisiko assoziiert sind. Im Zentrum der Netzwerkmodelle standen vier bekannte Tumorsuppressoren für Brustkrebs: BRCA1, BRCA2, ATM und CHEK2. Die angewandte Strategie kombinierte Genexpressionsprofile

mit funktionellen genomischen und proteomischen Daten von verschiedenen Spezies. Mit diesen Daten konnte ein Netzwerk um BRCA1, BRCA2, ATM und CHEK2 aus insgesamt 118 Genen erstellt werden, die durch mehrere hundert potenzielle funktionelle Assoziationen verknüpft sind. Pujana et al. (2007) [45] vermuteten, dass die Komponenten des Netzwerks in verwandten biologischen Signalwegen wirken und die 114 Gene um BRCA1, BRCA2, ATM und CHEK2 somit interessante Kandidatengene für Brustkrebsprädisposition darstellen. Neun Gene des komplexen Netzwerks liegen innerhalb der in dieser Arbeit identifizierten Kopplungsregionen: RAD51C (17q), CD1C (1q), MYBL2 (20q), UBE2C (20q), PAXIP1L (7q), TFDP1 (13q), DDX39 (19), ILF3 (19) und ZNF43 (19). Diese neun Gene sind in der Tabelle der 62 Kandidatengene (vgl. Abschnitt 3.2), sofern sie in der Kandidatengenanalyse nicht den Rang 1 oder 2 erhalten haben, als mögliche Kandidatengene für Brustkrebs susceptibility mit aufgeführt.

Die in dieser Arbeit angewandte Bewertung der Kandidatengene nach ihrem Rang (1 bis 4) und nach den Ergebnissen von Pujana et al. (2007) [45] ermöglichte die Gesamtzahl der 866 Gene in den acht Kopplungsregionen auf 62 Kandidatengene einzuschränken. Die Problematik der auf einer Datenbank- und Literaturrecherche basierenden Kandidatengenanalyse lag jedoch auch hierbei an der Abhängigkeit vom aktuellen Stand der Wissenschaft über die Funktion der Gene und an einer möglicherweise falsch erfolgten Rangzuordnung. Die Methode weist somit einige Fehlermöglichkeiten auf, die zu einer Identifikation von irrelevanten Kandidatengen bzw. zum Ausschluss des einzig richtigen Krankheitsgens geführt haben könnte. Eine weitere Verfolgung bzw. der Fortschritt in der Forschung und der Funktionsidentifikation bringt weitere Möglichkeiten der Detaillierung und Verbesserung der Ergebnisse mit sich.

Zuletzt wurden aus den Kandidatengen diejenigen ausgewählt, die durch Sequenzanalyse auf eine krankheitsverursachende Mutation analysiert werden sollten. Hierfür wurden vier Kriterien angesetzt, nach deren Übereinstimmung die Kandidatengene bewertet wurden (vgl. Abschnitt 3.2). Dies erfolgte ebenfalls durch Datenbank- und Literaturrecherche.

Das erste Kriterium war eine mögliche Rolle des Gens als Tumorsuppressor bzw. als Onkogen. Als klassische Tumorgene wird ihnen bei der Karzinogenese eine wichtige Rolle zugeschrieben [2]. Das nächste Kriterium berücksichtigte die Assoziation des Gens mit Brustkrebs, d.h., dass in Studien eine Expressionsveränderung, Amplifikation oder LOH des Gens mit Brustkrebs beschrieben worden war. Solche Veränderungen weisen auf eine mögliche Bedeutung des Gens während der Brustkrebsentstehung hin. Ferner zeigte die Betrachtung der Funktion verschiedener Suszeptibilitätsgene für Brustkrebs, dass ein kritischer biologischer Signalweg in der Entstehung dieser Krebsart die Antwort auf DNA-Schädigung ist. Die bekannten Brustkrebsgene BRCA1 und BRCA2 wirken als Sensoren von DNA-Schädigung und sind an Reparaturprozessen, u.a. der Reparatur von DNA-Doppelstrangbrüchen durch homologe Rekombination, beteiligt [21]. Die Suszeptibilitätsgene ATM, CHEK2, PALB2 und BRIP1 interagieren mit BRCA1 und/oder BRCA2 in diesen Prozessen. So wurde eine mögliche Funktion in der DNA-Reparatur mit dem dritten Kriterium bewertet. Das vierte Kriterium überprüfte die Funktion des Gens im Apoptose-Signalweg, in der Proliferation oder in der Zellzyklus-Kontrolle. Positive Wirkung des Gens auf die Proliferation bzw. negative Wirkung auf die Apoptose können die Karzinogenese beeinflussen.

Durch diese Zuordnung konnte eine Rangordnung innerhalb der Kandidatengene erreicht werden, die es erleichterte die Auswahl der "besten" Kandidatengene für die Sequenzie-



rung zu treffen. Fünf Gene erfüllten mindestens drei der aufgestellten Kriterien:

### ***BRIP1***

BRIP1 ist ein bekanntes Suszeptibilitätsgen für Brustkrebs mit mäßiger Penetranz (vgl. Abschnitt 1.2.3). Es stellte somit das "beste" Kandidatengen für den in der untersuchten Familie auftretenden Brustkrebs dar. Dieses Gen war bereits im Vorfeld der Arbeit durch Sequenzanalyse für Person III-9 untersucht worden. Es konnten jedoch keine pathogenen Veränderungen der DNA-Sequenz gefunden werden.

### ***RAD51C/RAD51L2***

Die homologe Rekombination ist ein wichtiger Mechanismus der Reparatur von DNA-Doppelstrangbrüchen. Die RAD51C-Funktion wird benötigt für normale Proliferation und spielt eine Rolle in der RAD51-vermittelten homologen Rekombination. Es wird vermutet, dass RAD51C mit der Lösung der *Holliday junctions*, dem letzten Schritt im Reparaturprozess der homologen Rekombination, assoziiert ist und auch eine Funktion in der frühen Phase dieser hat [47, 48]. Ferner ist RAD51C wichtig für die Aufrechterhaltung der korrekten Zentrosomenanzahl während der Mitose. Der Verlust einer korrekten Chromosomensegregation ist häufig mit der Entstehung von Krebs assoziiert. RAD51C verhindert vermutlich die Entstehung zellulärer Aneuploidie und ist ein potenzielles Tumorsuppressorgen [49].

BRCA1 und BRCA2, die ebenfalls in der DNA-Reparatur durch homologe Rekombination wirken, sind auch mit Aufrechterhaltung der Zentrosomenanzahl assoziiert. RAD51C wirkt in ähnlichen zellulären Prozessen wie BRCA1 und BRCA2 und stellt damit ein sehr interessantes Kandidatengen für Brustkrebs-suszeptibilität dar. Es entsprach in der Kandidatengenanalyse denselben drei Kriterien wie das BRIP1-Gen und wurde zusätzlich von Pujana et al. (2007) [45] als Komponente des komplexen Netzwerks um BRCA1, BRCA2, ATM und CHEK2 identifiziert.

### ***SLC9A3R1/NHERF1***

Das Gen SLC9A3R1 kodiert für ein Protein, das den epithelialen  $Na^+/H^+$ -Austauscher Isoform 3 inhibitorisch reguliert. Ferner interagiert das Genprodukt mit dem putativen Brustkrebs- oder Metastasensuppressor SYK und mit MERLIN, dem Produkt des Neurofibromatose 2 (NF2) Tumorsuppressorgens [50]. Mutationen im NHERF1-Gen begleitet von LOH des zweiten Allels sind in Brustkrebszelllinien und primären Mammakarzinomen gefunden worden. Diese Mutationen verhinderten oder schwächten eine Interaktion mit SYK oder MERLIN ab [51]. Die biologische Aktivität von NHERF1 steht in der Brustdrüse vermutlich mit dem PTEN-Signalweg in Verbindung [52]. NHERF1 könnte ein mögliches Tumorsuppressorgen für Brustkrebs sein, das mit dem SYK- und MERLIN-Suppressor verbunden ist. Die Tumorsuppressoraktivität in der Brust könnte auch durch einen regulatorischen Effekt auf die Zellzyklusprogression begründet sein [53]. SLC9A3R1 erfüllte drei Kriterien der Kandidatengenanalyse. Es stellt ein mögliches Kandidatengen für einen Tumorsuppressor in der Brust dar.

### ***PYHIN1/IFIX***

PYHIN1 kodiert für ein Protein der HIN-200 Proteinfamilie, die durch Interferon induziert wird. Die Expression des Gens ist in vielen Mammakarzinomen und Brustkrebszelllinien reduziert. Untersuchungen der längsten Isoform IFIXa1 zeigten, dass dessen Expression mit Wachstumsinhibition, Suppression der Transformation und Tumorgenese assoziiert ist. IFIXa1 fungiert als Tumorsuppressor [54, 55]. Es entsprach drei Kriterien

in der Kandidatengenanalyse.

### **ING1**

ING1-Proteine wirken in verschiedenen zellulären Prozessen, einschließlich der Unterdrückung von Zellwachstum und Proliferation, Zellzyklusarrest, DNA-Reparatur, Apoptose und der Aufrechterhaltung der genomischen Stabilität. Sie sind assoziiert mit TP53 in der Aktivierung von p21, dessen Genprodukt den p53-abhängigen Zellzyklusarrest vermittelt. Ferner haben ING1-Proteine eine wichtige Rolle in der transkriptionalen Regulation der Expression verschiedener Gene. Sie können die Transkription des proapoptotischen Bax-Gens aktivieren und damit die Apoptose fördern. Eine Funktion ist weiter die Interaktion mit DNA-Reparaturfaktoren, die der Beseitigung von UV-induzierten DNA-Läsionen dient. Isoform ING1b ist ein wichtiger Faktor der Nukleotid-Exzisionsreparatur [56, 57]. Im Mammakarzinom wird selten eine Alteration des ING1-Gens, jedoch häufig eine Herunterregulierung der nukleären Expression beobachtet. ING1 ist ein Tumorsuppressorgen, dessen Verlust oder Inaktivierung zu veränderter Zellwachstumskontrolle, Apoptoseresistenz und sogar einem immortalen Tumorphänotyp führen kann [58]. Dieses Gen stellt ebenfalls ein interessantes Kandidatengen für Brustkrebsprädisposition dar.

Die Sequenzanalyse ist eine Methode um in Kandidatengen nach patientenspezifischen Mutationen zu suchen. Ein Kandidatengen ist an der Pathogenese einer Erkrankung beteiligt, wenn es bei erkrankten Individuen eine Mutation trägt, die bei nicht betroffenen Individuen nicht vorkommt. Eine Mutation, die nur bei betroffenen Frauen der Familie zu finden ist, würde darauf hinweisen, dass dieses Gen wahrscheinlich der Krankheitslokus ist. Um jedoch letzten Endes die Authentizität des Kandidatengens begründen zu können, muss das mutierte Allel nicht nur mit der Krebsprädisposition segregieren, sondern auch zur Entstehung der bösartigen Neubildung beitragen. Es muss geklärt werden, inwiefern die gefundene Sequenzveränderung pathogen ist. Dafür sind funktionelle Studien durchzuführen.

Die Methode der Sequenzierung ermöglicht die Detektion von Punktmutationen, kurzen Deletionen, Insertionen, Inversionen und Frame-Shift-Mutationen. Mit Ausschluss einer Mutation innerhalb der Exone der untersuchten Kandidatengene ist die Beteiligung des jeweiligen Gens an der Auslösung eines Mammakarzinoms dennoch nicht vollkommen ausgeschlossen. Liegt ein großes genomisches Rearrangement vor, ist die Grenze dieser Methode erreicht. Bei heterozygoter Deletion ganzer Exone zeigt sich in der Sequenzanalyse die Basenabfolge des gesunden Allels, die keinen Hinweis auf eine Mutation trägt. Ebenfalls schwer erkennbar ist die Inversion eines ganzen Exons, wenn die intronischen Regionen, an die das Primerpaar bindet, ebenfalls von dieser betroffen sind. Ist solch ein Rearrangement in der untersuchten Familie verantwortlich für die Brustkrebsprädisposition, konnte der Krankheitslokus mit der angewandten Methode nicht gefunden werden.

Die untersuchten kodierenden Exone der Kandidatengene RAD51C, PYHIN1, ING1 und SLC9A3R1 wurden als vollständig ausgewertet betrachtet, wenn für wenigstens eine betroffene Person und für die Kontrollperson die DNA-Sequenz mindestens 20 bp vor und hinter dem Exon-Intron-Übergang entweder durch den Forward- oder den Reverse-Primer lesbar war. Dieser minimal analysierte intronische Bereich deckt die bekannten Spleißakzeptor- und Spleißdonor-Konsensus-Sequenzen ab. Durch das Spleißen werden die Introne entfernt und die angrenzenden Exone miteinander zu einer fertigen mRNA verknüpft. Mutationen in einem Intron haben keinen direkten Effekt auf die Sequenz

des Proteins, das durch ein Gen kodiert wird. Mutagenese kann jedoch überall im Intron kryptische Spleißstellen aktivieren, die zu einem modifizierten Prozessieren der prä-mRNA führen. Die so entstehenden RNAs kodieren für veränderte Proteine und können ursächlich für eine Erkrankung sein [3]. Die Bewertung einer intronisch gefundenen Mutation bereitet somit zusätzliche Probleme. Da lediglich die Exone der Gene sequenziert wurden, kann über mögliche intronische Mutationen keine Aussage gemacht werden.

Zusätzlich ist die Transkription vieler Gene sehr streng reguliert. So verfügen die Gene über einen meist am 5'-Ende der kodierenden Sequenz gelegenen Abschnitt, der die benötigte Basensequenz für die Bindung des Initiationskomplexes der Transkription sowie für zusätzliche regulatorische Transkriptionsfaktoren enthält. Dieser wird als Promotor bezeichnet. Die spezifischen Bereiche der Promotorregion werden TATA-Box (-20 bp), GC-Box (-60 bp und -100 bp), CAAT-Box (-80 bp) und Oktamer (-120 bp) genannt [3]. Mutationen in diesen Bereichen wurden in dieser Arbeit nicht explizit untersucht.

Da die Analyse des Stammbaums der untersuchten Familie einen autosomal dominant vererbten genetischen Defekt vermuten lässt, kann man davon ausgehen, dass alle Betroffenen und Überträger des Gendefekts die verursachende Mutation in der untersuchten DNA peripherer Blutzellen nur in einer ihrer homologen DNA-Sequenzen aufweisen. Gesunde Kontrollpersonen sollten diese Mutation nicht tragen. Es konnte für die sequenzierten Exone keine Mutation im Vergleich zu einem gesunden Individuum und der Gensequenz der Genomsequenz hg18 der UCSC Genome Bioinformatics Datenbank gefunden werden, die dieses Kriterium erfüllt. Bei den gefundenen vier Sequenzveränderungen (vgl. Abschnitt 3.3) handelt es sich jeweils um bekannte polymorphe SNPs. Für das BRIP1 Gen waren bereits im Vorfeld der Arbeit die 20 Exone inklusive der angrenzenden intronischen Bereiche für Person III-9 sequenziert worden. Es konnte ebenfalls keine pathogene Veränderung nachgewiesen werden. Es fanden sich lediglich fünf häufig vorkommende Polymorphismen in Intron 8 (IVS891\_92delT), Intron 9 (IVS91109G>A), Intron 13 (IVS13-47G>C), Exon 19 (c.2757T>C) und Exon 20 (c.4501T>C), jeweils heterozygot.

Der genetische Defekt, der in der Familie ursächlich für den familiär auftretenden Brustkrebs ist, konnte im Rahmen dieser Arbeit nicht identifiziert werden. Es kann jedoch nicht ausgeschlossen werden, dass für BRIP1, RAD51C, PPH1, ING1 und SLC9A3R1 die krankheitsverursachende Mutation mit der Methode der Sequenzierung nicht detektiert werden konnte. Sie stellen weiterhin Kandidatengene für den in der Familie auftretenden Brustkrebs dar.

Familienbasierte Kopplungsstudien sind sehr erfolgreich in der Identifizierung von Genen, die einer monogenetisch verursachten Erkrankung zugrunde liegen. Jedoch waren Versuche weitere Gene zu identifizieren, die mit einem erhöhten Brustkrebsrisiko ähnlich BRCA1 oder BRCA2 einhergehen, bis heute erfolglos. Im Folgenden werden die in dieser Arbeit gefundenen acht Kandidatenregionen für Brustkrebs susceptibility mit denen anderer Studien verglichen.

Verlust der Heterozygotität- (*loss of heterozygosity*, LOH) und Kopplungsdaten ließen Kerangueven et al. (1995) [59] die chromosomale Region 8p12-p22 als Locus vermuten, der Tumorsuppressoren enthält. Seitz et al. (1997) [60] beobachtete für zwei deutsche Brustkrebsfamilien ein positives Kopplungssignal für diese Region. Bei Kainu et al. (2000) [61] wurde unter Anwendung vergleichender genomischer Hybridisierung (*comparative genomic hybridization*, CGH) 13q21 als möglicher Brustkrebs susceptibility locus

propagiert. Keiner dieser Loki wurde von Rahman et al. (2000) [62] bzw. Thompson et al. (2002) [63] bestätigt. Auch in der vorliegenden Arbeit gehören die Regionen 8p12-p22 und 13q21 nicht zu den Kandidatenregionen.

Huusko et al. (2004) [64] identifizierten mit Hilfe einer genomweiten Kopplungsanalyse, die 14 finnische Brustkrebsfamilien negativ für eine Mutation in den Genen BRCA1 oder BRCA2 einschloss, einen möglichen Suszeptibilitätsloкус in der chromosomalen Region 2q24-q33 (NPLD 3,20) und einen weiteren in 9q21 (LOD 1,12). In der Studie von Bergman et al. (2007) [65] zeigte die Analyse von 14 schwedischen nicht-BRCA1/2-Familien (1 große und 13 kleine bis mittelgroße Brustkrebsfamilien) die chromosomale Region 10q23.32-q25.3 (HLOD 2,34) als wahrscheinlichsten Lokus für eine Kopplung an. Zwei weitere Regionen wurden bei 12q14-q21 (HLOD 2,16) und 19p13.3-q12 (HLOD 2,10) identifiziert. Die große Brustkrebsfamilie allein erreichte einen LOD  $> 1,5$  für die schon beschriebenen Kandidatenregionen auf 10q23.32-q25.3 (HLOD 1,66) und 19q13.12-q13.32 (HLOD 1,52), sowie einen weiteren für die Region 17p13 (HLOD 1,51). Bei Ausschluss der großen Familie aus der Berechnung wurde das wahrscheinlichste Kopplungssignal am 19p13-q12 Lokus (HLOD 2,10) erreicht, gefolgt von einem HLOD von 1,91 bei 16p13.2-q12.1. Der Kandidatenloкус 19p13.3-q12 für die 14 schwedischen Familien stimmt mit der zweit-wahrscheinlichsten Region für Kopplung bei der in dieser Arbeit untersuchten Familie überein. Diese erreichte in der parametrischen Kopplungsanalyse für die chromosomale Region 19p13.2-q12 einen maximalen multipoint LOD-Wert von 2,444. Der genaue Überlappungsbereich kann nicht angegeben werden, da Bergman et al. (2007) [65] die exakten Grenzen der Kopplungsregionen nicht bestimmten. Das Ergebnis lässt jedoch vermuten, dass die Region 19p13.3-q12 ein Suszeptibilitätsgen für Brustkrebs enthält, das in diesen Familien für die Entstehung der Erkrankung verantwortlich sein kann. Diese Überlappung könnte ein interessanter Angriffspunkt für weitere Arbeiten darstellen. Die anderen Kandidatenregionen der finnischen und schwedischen Studie konnten in dieser Arbeit nicht bestätigt werden. Sie können jedoch spezifisch entweder für die finnische oder die schwedische Population sein.

Smith et al. (2006) [66] führten eine kombinierte Analyse von genomweiten Kopplungsuntersuchungen mit einer Gesamtzahl von 149 Brustkrebsfamilien durch. Es wurden nur Familien kaukasischen Ursprungs mit mindestens drei Brustkrebsdiagnosen unter einem Alter von 60 Jahren in die Studie miteinbezogen. Wenigstens eine Erkrankte musste im Vorfeld negativ auf pathogene Mutationen in den Genen BRCA1 und BRCA2 getestet worden sein. Der höchste HLOD-Wert von 1,80 wurde in der parametrischen Analyse unter Annahme eines dominanten Modells für die chromosomale Region 4q erreicht. Weitere Regionen mit einem LOD  $> 1$  wurden auf den Chromosomen 2p (HLOD 1,21; NPLD 1,10), 5 (HLOD 1,04), 14 (NPLD 1,56) und 22 (HLOD 1,15) identifiziert. Bei Einschränkung der Analyse auf diejenigen Familien mit mindestens vier Brustkrebserkrankungen, die alle unter einem Alter von 50 Jahren diagnostiziert wurden, zeigte sich die wahrscheinlichste Kandidatenregion auf 2p (HLOD 2,38). Ein HLOD  $> 1$  wurde in dieser Gruppe auf drei weiteren Chromosomen erreicht: 4 (HLOD 1,57), 10 (HLOD 1,12) und 22 (HLOD 1,43). Obwohl die Anzahl der analysierten Familien sehr groß war, konnte keine Region identifiziert werden, die dem klassischen Signifikanzkriterium von einem LOD-Wert  $> 3,0$  entsprach. Das Fehlen eines signifikanten Kopplungssignals ist in dieser Studie vermutlich auf Heterogenität der untersuchten Familien zurückzuführen. Um additionalen Brustkrebs-suszeptibilitätsgene zu detektieren, scheint es essenziell, die genetische Heterogenität zu reduzieren. Keine in dieser Arbeit gefundene Kandidatenregion überlappt mit denen in der Analyse von Smith et al. (2006) [66] gefundenen Regionen.

Jedoch wurde durch die Analyse von nur einer Familie das Problem der genetischen Heterogenität wahrscheinlich umgangen.

Gonzalez-Neira et al. (2007) [46] führten eine genomweite Kopplungsanalyse in 19 nicht-BRCA1/2-Brustkrebsfamilien durch. Sie identifizierten fünf mögliche Kandidatenregionen auf den Chromosomenabschnitten 2p22.2 (NPLD 2,26), 4p14-q12 (NPLD 2,29), 7q21.11-q21.3 (NPLD 2,56), 11q13.5-q14.3 (NPLD 2,21) und 14q21.1-q21.3 (NPLD 2,25). Auch diese möglichen Kandidatenregionen für Brustkrebs-suszeptibilität konnten in der hier durchgeführten Kopplungsanalyse nicht bestätigt werden.

Die Regionen auf den Chromosomen Xq, 17q, 1q, 20q, 7q, 8p und 13q wurden in dieser Arbeit erstmals als Kandidatenregionen für Brustkrebs-suszeptibilität beschrieben.

Die Gene BRCA1 und BRCA2 sind für 16% des familiär bedingten Risikos für Brustkrebs verantwortlich [14]. Mutationen in diesen beiden Genen waren bereits im Vorfeld der Arbeit durch DHPLC-Analyse und direkte Sequenzierung für die Personen III-5, III-6 und III-9 ausgeschlossen worden. Die chromosomalen Regionen des BRCA1- und BRCA2-Gens sollten damit auch keinen Hinweis auf Kopplung in der durchgeführten Analyse zeigen. Ein positiver LOD deutet auf Kopplung hin, ein negativer Wert dagegen auf Abwesenheit von Kopplung. Bei einem LOD-Wert  $< -2$  kann die Kopplung der Marker mit dem Krankheitsloкус signifikant ausgeschlossen werden [3]. Die Überprüfung des BRCA1-Lokus in der Output-Datei der parametrischen Merlin-Analyse zeigte, dass die zugehörige chromosomale Region in den durchgeführten Kopplungsanalysen mit Individuum III-5 und III-6 "unknown" einen negativen mLOD-Wert  $< -2$  besaß. Für das BRCA1-Gen wurde bei der Analyse mit Individuum III-1 auf "unknown" gesetzt ein maximaler mLOD von -0,923 erreicht (III-7 "unknown":  $\text{mLOD} \leq -1,845$ ; III-9 "unknown":  $\text{mLOD} \leq -1,823$ ; III-10 "unknown":  $\text{mLOD} \leq -1,837$ ). Für den BRCA2-Lokus konnte mit einem mLOD-Wert  $< -2$  in allen Analysen eine Kopplung definitiv ausgeschlossen werden. Ein Fehlen jeglicher Kopplung zu BRCA1 und BRCA2 war somit in der dieser Arbeit zugrunde liegenden Analyse nachweisbar.

#### 4.1 Bilanz und Ausblick

Im Rahmen dieser Arbeit wurde eine genomweite Kopplungsanalyse für das autosomal dominant vererbte Mammakarzinom durchgeführt. Durch die Analyse einer Brustkrebs-Familie konnten acht interessante Kopplungsregionen bestimmt werden. Keine dieser Regionen erzielte einen signifikanten LOD-Wert, was jedoch auf die Größe und Struktur der untersuchten Familie zurückzuführen ist. Die Kandidatenregion 19p13.2-q12 war bereits in der Analyse von Bergman et al. (2007) [65] gefunden worden. Die chromosomalen Bereiche Xq27.2-q28, 17q21.33-q25.1, 1q23.1-q23.3, 20q12-q13.2, 7q36.2-q36.3, 8p23.3 und 13q32.1-q34 wurden in dieser Arbeit erstmals als Kandidatenregionen für Brustkrebs-suszeptibilität beschrieben. Im Anschluss an die Identifikation der Kopplungsregionen wurde in den acht Kandidatenregionen nach möglichen Kandidatengenen für die Erkrankung Brustkrebs gesucht. Fünf der 62 identifizierten möglichen Kandidatengene (BRIP1 (17q), RAD51C (17q), SLC9A3R1 (17q), PYHIN1 (1q) und ING1 (13q)) wurden durch eine Sequenzierung der kodierenden Exone analysiert. Es konnte jedoch keine krankheits-assoziierte Mutation aufgedeckt werden.

Ausgehend von den vorliegenden Ergebnissen stellt sich zunächst die Frage nach der Richtigkeit der gefundenen Kopplungsregionen. Durch Einbeziehung weiterer Mitglieder der Familie aus fünfter Generation könnten falsch positive Kandidatenregionen ausge-

geschlossen werden. Bis in dieser Generation jedoch eine Diagnose Brustkrebs gestellt wird, können noch einige Jahre vergehen. Eine andere Möglichkeit ist, die weiteren 57 Kandidatengene einer Mutationssuche zu unterziehen. Es muss auch weiterhin beobachtet werden, ob neue Kandidatengene innerhalb der hier ermittelten Kopplungsregionen gefunden werden. Hierfür ist es unentbehrlich in regelmäßigen zeitlichen Abständen die Gendatenbanken erneut zu analysieren, um zwischenzeitlich identifizierte Gene in die Liste der Kandidatengene aufnehmen zu können. Schließlich sollten weitere Methoden der Identifikation einer vorhandenen Mutation, die möglicherweise außerhalb der in dieser Arbeit sequenzierten Exone liegt, angestrebt werden.

Obwohl neu gefundene Suszeptibilitätsgene für Brustkrebs wahrscheinlich nur für eine kleine Fraktion der nicht-BRCA1/2-Brustkrebsfamilien verantwortlich sein werden, ist deren Identifizierung von großem wissenschaftlichen Wert um die zellulären Vorgänge, die zur Entstehung von Brustkrebs führen, besser zu verstehen. Das Wissen ist auch vor allem für die Trägerfamilien von großer Bedeutung, da die Identifikation eines neuen Suszeptibilitätsgens ermöglichen würde Risikopatientinnen in engmaschige Vorsorgeprogramme zu integrieren und ihnen frühzeitig die bestmöglichen Therapie- und Behandlungsprogramme anzubieten.

## 5 ZUSAMMENFASSUNG

In der vorliegenden Arbeit standen Blutproben einer großen Brustkrebs-Familie mit mehreren betroffenen Familienmitgliedern, die negativ auf Mutationen in den beiden Brustkrebsgenen BRCA1 und BRCA2 getestet wurden, zur Verfügung. Damit wurde eine genomweite Kopplungsanalyse mit einem anschließenden positionellen Kandidatengenansatz zur Auffindung der genetischen Ursache des Mammakarzinoms in dieser Familie durchgeführt. Die elf zur Verfügung stehenden DNA Proben wurden mit dem GeneChip Human Mapping 50K Array XbaI 240 der Firma Affymetrix genotypisiert. Insgesamt wurden 35.642 Einzelnukleotidpolymorphismen (SNPs) über alle Chromosomen verteilt in die Kopplungsanalyse einbezogen.

Die Stammbaumanalyse der Familie lässt vermuten, dass der zugrunde liegende genetische Defekt für das Mammakarzinom dominant vererbt wird. Entsprechend wurde die parametrische Kopplungsanalyse unter dieser Annahme, jedoch auch unter nicht-parametrischen Modellen, durchgeführt. Unter Zugrundelegung eines dominanten Defekts zeigten sich die höchsten Kopplungssignale in den chromosomalen Regionen Xq (mLOD 1,8) und 17q (mLOD 1,35). Bei Berücksichtigung von sporadischen Phänokopien in der Familie konnten sieben Kopplungsregionen mit einem LOD-Wert  $> 1,4$  auf den Autosomen identifiziert werden. Diese lagen auf den Chromosomen 1q (mLOD 1,949), 17q (mLOD 1,819), 20q (mLOD 2,066), 7q (mLOD 1,465), 8p (mLOD 2,654), 13q (mLOD 2,429) und 19 (mLOD 2,444). Die in der parametrischen Kopplungsanalyse identifizierten acht Kandidatenregionen zeigten jeweils auch in der nicht-parametrischen Analyse die besten Ergebnisse. Keine der möglichen Kopplungsregionen erzielte einen signifikanten LOD-Wert, was jedoch aufgrund der Struktur und Größe der untersuchten Familie und der damit eingeschränkten statistischen Aussagekraft nicht erwartet werden konnte.

Im Rahmen einer Haplotypenanalyse wurden die segregierenden krankheitsspezifischen Haplotypen identifiziert. Diese wurden anschließend durch die Genotypisierung von Mikrosatelliten-Markern bestätigt und in einigen Fällen weiter eingegrenzt. Die chromosomalen Bereiche Xq27.2-q28, 17q21.33-q25.1, 1q23.1-q23.3, 20q12-q13.2, 7q36.2-q36.3, 8p23.3 und 13q32.1-q34 wurden in dieser Arbeit erstmals als Kandidatenregionen für Brustkrebs-suszeptibilität beschrieben. Die Kandidatenregion 19p13.2-q12 war bereits in der Analyse von Bergman et al. (2007) [65] aufgefunden worden.

Nach der Identifizierung der Kopplungsregionen wurde in den jeweiligen chromosomalen Bereichen nach möglichen Kandidatengenen für Brustkrebs gesucht. Die Gesamtzahl von 866 Genen in den acht Kopplungsregionen konnte durch eine Reihe bioinformatischer Verfahren auf 62 Kandidatengene eingeschränkt werden. Fünf der 62 identifizierten Kandidatengene (BRIP1, RAD51C, PYHIN1, ING1 und SLC9A3R1) wurden durch eine Sequenzierung der kodierenden Exone analysiert. Es konnte jedoch keine krankheitsassoziierte Mutation gefunden werden.

Ungeachtet der systematischen Vorgehensweise konnte der genetische Defekt, der in der untersuchten Familie für das familiäre Auftreten von Brustkrebs verantwortlich ist, im Rahmen dieser Arbeit nicht identifiziert werden. Dies zeigt unter anderem die häufige Problematik des positionellen Kandidatengenansatzes bei Krankheiten mit komplexer Ätiologie, der sich in seiner Analyse auf einzelne (kleine) Familien beschränken muß.

Die Arbeit wurde in dem Zeitraum zwischen Oktober 2007 und April 2008 am Institut für Humangenetik der Universität Regensburg durchgeführt.

## 6 LITERATURVERZEICHNIS

- [1] A. Balmain, J. Gray, and B. Ponder. The genetics and genomics of cancer. *Nat. Genet.*, 33 Suppl:238–244, Mar 2003.
- [2] W. Böcker, H. Denk, and P.U. Heitz. Pathologie. *Urban Fischer Verlag/Elsevier GmbH*, 3, 2004.
- [3] T. Strachan and A. P. Read. Molekulare Humangenetik. *Spektrum Akademischer Verlag*, 3, 2005.
- [4] E. R. Fearon. Human cancer syndromes: clues to the origin and nature of cancer. *Science*, 278(5340):1043–1050, Nov 1997.
- [5] A. G. Knudson. Mutation and cancer: statistical study of retinoblastoma. *Proc. Natl. Acad. Sci. U.S.A.*, 68(4):820–823, Apr 1971.
- [6] A. G. Knudson. Two genetic hits (more or less) to cancer. *Nat. Rev. Cancer*, 1(2):157–162, Nov 2001.
- [7] Robert Koch-Institut. Gesundheit in Deutschland. Gesundheitsberichterstattung des Bundes. *Robert Koch-Institut, Berlin*, <http://www.gbe-bund.de>, 2006.
- [8] Robert Koch-Institut und Gesellschaft der epidemiologischen Krebsregister in Deutschland e. V. Krebs in Deutschland 2003-2004. Häufigkeiten und Trends. *Robert Koch-Institut, Berlin*, 6, 2008.
- [9] Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and breastfeeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50302 women with breast cancer and 96973 women without the disease. *Lancet*, 360:187–195, Jul 2002.
- [10] Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52,705 women with breast cancer and 108,411 women without breast cancer. *Lancet*, 350:1047–1059, Oct 1997.
- [11] L. Bernstein. The risk of breast, endometrial and ovarian cancer in users of hormonal preparations. *Basic Clin. Pharmacol. Toxicol.*, 98:288–296, Mar 2006.
- [12] M. R. Stratton and N. Rahman. The emerging landscape of breast cancer susceptibility. *Nat. Genet.*, 40(1):17–22, Jan 2008.
- [13] C. Szpirer and J. Szpirer. Mammary cancer susceptibility: human genes and rodent models. *Mamm. Genome*, 18(12):817–831, Dec 2007.
- [14] R. A. Oldenburg, H. Meijers-Heijboer, C. J. Cornelisse, and P. Devilee. Genetic susceptibility for breast cancer: how many more genes to be found? *Crit. Rev. Oncol. Hematol.*, 63(2):125–149, Aug 2007.
- [15] J. M. Hall, M. K. Lee, B. Newman, J. E. Morrow, L. A. Anderson, B. Huey, and M. C. King. Linkage of early-onset familial breast cancer to chromosome 17q21. *Science*, 250(4988):1684–1689, Dec 1990.



- [16] Y. Miki, J. Swensen, D. Shattuck-Eidens, P. A. Futreal, K. Harshman, S. Tavtigian, Q. Liu, C. Cochran, L. M. Bennett, and W. Ding. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science*, 266(5182):66–71, Oct 1994.
- [17] R. Wooster, S. L. Neuhausen, J. Mangion, Y. Quirk, D. Ford, N. Collins, K. Nguyen, S. Seal, T. Tran, and D. Averill. Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13. *Science*, 265(5181):2088–2090, Sep 1994.
- [18] R. Wooster, G. Bignell, J. Lancaster, S. Swift, S. Seal, J. Mangion, N. Collins, S. Gregory, C. Gumbs, and G. Micklem. Identification of the breast cancer susceptibility gene BRCA2. *Nature*, 378(6559):789–792, 1995.
- [19] K. L. Nathanson, R. Wooster, B. L. Weber, and K. N. Nathanson. Breast cancer genetics: what we know and what we need. *Nat. Med.*, 7(5):552–556, May 2001.
- [20] D. Thompson and D. Easton. The genetic epidemiology of breast cancer genes. *J Mammary Gland Biol Neoplasia*, 9(3):221–236, Jul 2004.
- [21] K. Gudmundsdottir and A. Ashworth. The roles of BRCA1 and BRCA2 and associated proteins in the maintenance of genomic stability. *Oncogene*, 25(43):5864–5874, Sep 2006.
- [22] N. G. Howlett, T. Taniguchi, S. Olson, B. Cox, Q. Waisfisz, C. De Die-Smulders, N. Persky, M. Grompe, H. Joenje, G. Pals, H. Ikeda, E. A. Fox, and A. D. D’Andrea. Biallelic inactivation of BRCA2 in Fanconi anemia. *Science*, 297(5581):606–609, Jul 2002.
- [23] A. C. Antoniou, P. D. Pharoah, G. McMullan, N. E. Day, M. R. Stratton, J. Peto, B. J. Ponder, and D. F. Easton. A comprehensive model for familial breast cancer incorporating BRCA1, BRCA2 and other genes. *Br. J. Cancer*, 86(1):76–83, Jan 2002.
- [24] A. R. Bradbury and O. I. Olopade. Genetic susceptibility to breast cancer. *Rev Endocr Metab Disord*, 8(3):255–267, Sep 2007.
- [25] H. Meijers-Heijboer, A. van den Ouweland, J. Klijn, M. Wasielewski, A. de Snoo, R. Oldenburg, A. Hollestelle, M. Houben, E. Crepin, M. van Veghel-Plandsoen, F. Elstrodt, C. van Duijn, C. Bartels, C. Meijers, M. Schutte, L. McGuffog, D. Thompson, D. Easton, N. Sodha, S. Seal, R. Barfoot, J. Mangion, J. Chang-Claude, D. Eccles, R. Eeles, D. G. Evans, R. Houlston, V. Murday, S. Narod, T. Peretz, J. Peto, C. Phelan, H. X. Zhang, C. Szabo, P. Devilee, D. Goldgar, P. A. Futreal, K. L. Nathanson, B. Weber, N. Rahman, and M. R. Stratton. Low-penetrance susceptibility to breast cancer due to CHEK2(\*)1100delC in noncarriers of BRCA1 or BRCA2 mutations. *Nat. Genet.*, 31(1):55–59, May 2002.
- [26] N. Rahman, S. Seal, D. Thompson, P. Kelly, A. Renwick, A. Elliott, S. Reid, K. Spanova, R. Barfoot, T. Chagtai, H. Jayatilake, L. McGuffog, S. Hanks, D. G. Evans, D. Eccles, D. F. Easton, and M. R. Stratton. PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. *Nat. Genet.*, 39(2):165–167, Feb 2007.
- [27] K. J. Patel. Fanconi anemia and breast cancer susceptibility. *Nat. Genet.*, 39(2):142–143, Feb 2007.

- [28] S. Seal, D. Thompson, A. Renwick, A. Elliott, P. Kelly, R. Barfoot, T. Chagtai, H. Jayatilake, M. Ahmed, K. Spanova, B. North, L. McGuffog, D. G. Evans, D. Eccles, D. F. Easton, M. R. Stratton, and N. Rahman. Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles. *Nat. Genet.*, 38(11):1239–1241, Nov 2006.
- [29] D. F. Easton, K. A. Pooley, A. M. Dunning, P. D. Pharoah, D. Thompson, D. G. Ballinger, J. P. Struwing, J. Morrison, H. Field, R. Luben, N. Wareham, S. Ahmed, C. S. Healey, R. Bowman, K. B. Meyer, C. A. Haiman, L. K. Kolonel, B. E. Henderson, L. Le Marchand, P. Brennan, S. Sangrajrang, V. Gaborieau, F. Odefrey, C. Y. Shen, P. E. Wu, H. C. Wang, D. Eccles, D. G. Evans, J. Peto, O. Fletcher, N. Johnson, S. Seal, M. R. Stratton, N. Rahman, G. Chenevix-Trench, S. E. Bojesen, B. G. Nordestgaard, C. K. Axelsson, M. Garcia-Closas, L. Brinton, S. Chanock, J. Lissowska, B. Peplonska, H. Nevanlinna, R. Fagerholm, H. Eerola, D. Kang, K. Y. Yoo, D. Y. Noh, S. H. Ahn, D. J. Hunter, S. E. Hankinson, D. G. Cox, P. Hall, S. Wedren, J. Liu, Y. L. Low, N. Bogdanova, P. Schurmann, T. Dork, R. A. Tollenaar, C. E. Jacobi, P. Devilee, J. G. Klijn, A. J. Sigurdson, M. M. Doody, B. H. Alexander, J. Zhang, A. Cox, I. W. Brock, G. MacPherson, M. W. Reed, F. J. Couch, E. L. Goode, J. E. Olson, H. Meijers-Heijboer, A. van den Ouweland, A. Uitterlinden, F. Rivadeneira, R. L. Milne, G. Ribas, A. Gonzalez-Neira, J. Benitez, J. L. Hopper, M. McCredie, M. Southey, G. G. Giles, C. Schroen, C. Justenhoven, H. Brauch, U. Hamann, Y. D. Ko, A. B. Spurdle, J. Beesley, X. Chen, A. Mannermaa, V. M. Kosma, V. Kataja, J. Hartikainen, N. E. Day, D. R. Cox, and B. A. Ponder. Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature*, 447(7148):1087–1093, Jun 2007.
- [30] D. J. Hunter, P. Kraft, K. B. Jacobs, D. G. Cox, M. Yeager, S. E. Hankinson, S. Wacholder, Z. Wang, R. Welch, A. Hutchinson, J. Wang, K. Yu, N. Chatterjee, N. Orr, W. C. Willett, G. A. Colditz, R. G. Ziegler, C. D. Berg, S. S. Buys, C. A. McCarty, H. S. Feigelson, E. E. Calle, M. J. Thun, R. B. Hayes, M. Tucker, D. S. Gerhard, J. F. Fraumeni, R. N. Hoover, G. Thomas, and S. J. Chanock. A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat. Genet.*, 39(7):870–874, Jul 2007.
- [31] S. N. Stacey, A. Manolescu, P. Sulem, T. Rafnar, J. Gudmundsson, S. A. Gudjonsson, G. Masson, M. Jakobsdottir, S. Thorlacius, A. Helgason, K. K. Aben, L. J. Strobbe, M. T. Albers-Akkers, D. W. Swinkels, B. E. Henderson, L. N. Kolonel, L. Le Marchand, E. Millastre, R. Andres, J. Godino, M. D. Garcia-Prats, E. Polo, A. Tres, M. Mouy, J. Saemundsdottir, V. M. Backman, L. Gudmundsson, K. Kristjansson, J. T. Bergthorsson, J. Kostic, M. L. Frigge, F. Geller, D. Gudbjartsson, H. Sigurdsson, T. Jonsdottir, J. Hrafinkelsson, J. Johannsson, T. Sveinsson, G. Myrdal, H. N. Grimsson, T. Jonsson, S. von Holst, B. Werelius, S. Margolin, A. Lindblom, J. I. Mayordomo, C. A. Haiman, L. A. Kiemeny, O. T. Johannsson, J. R. Gulcher, U. Thorsteinsdottir, A. Kong, and K. Stefansson. Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat. Genet.*, 39(7):865–869, Jul 2007.
- [32] A. Cox, A. M. Dunning, M. Garcia-Closas, S. Balasubramanian, M. W. Reed, K. A. Pooley, S. Scollen, C. Baynes, B. A. Ponder, S. Chanock, J. Lissowska, L. Brinton, B. Peplonska, M. C. Southey, J. L. Hopper, M. R. McCredie, G. G. Giles, O. Fletcher, N. Johnson, I. dos Santos Silva, L. Gibson, S. E. Bojesen, B. G. Nordestgaard,

- C. K. Axelsson, D. Torres, U. Hamann, C. Justenhoven, H. Brauch, J. Chang-Claude, S. Kropp, A. Risch, S. Wang-Gohrke, P. Schurmann, N. Bogdanova, T. Dork, R. Fagerholm, K. Aaltonen, C. Blomqvist, H. Nevanlinna, S. Seal, A. Renwick, M. R. Stratton, N. Rahman, S. Sangrajrang, D. Hughes, F. Odefrey, P. Brennan, A. B. Spurdle, G. Chenevix-Trench, J. Beesley, A. Mannermaa, J. Hartikainen, V. Kataja, V. M. Kosma, F. J. Couch, J. E. Olson, E. L. Goode, A. Broeks, M. K. Schmidt, F. B. Hogervorst, L. J. Van't Veer, D. Kang, K. Y. Yoo, D. Y. Noh, S. H. Ahn, S. Wedren, P. Hall, Y. L. Low, J. Liu, R. L. Milne, G. Ribas, A. Gonzalez-Neira, J. Benitez, A. J. Sigurdson, D. L. Stredrick, B. H. Alexander, J. P. Struewing, P. D. Pharoah, and D. F. Easton. A common coding variant in CASP8 is associated with breast cancer risk. *Nat. Genet.*, 39(3):352–358, Mar 2007.
- [33] W. Buselmaier and G. Tariverdian. Humangenetik. *Springer Medizin Verlag Heidelberg*, 4, 2007.
- [34] J. Ott. Analysis of Human Genetic Linkage. *The Johns Hopkins University Press Baltimore, Maryland*, 1991.
- [35] N. E. MORTON. Sequential tests for the detection of linkage. *Am. J. Hum. Genet.*, 7(3):277–318, Sep 1955.
- [36] G. M. Lathrop, J. M. Lalouel, C. Julier, and J. Ott. Strategies for multilocus linkage analysis in humans. *Proc. Natl. Acad. Sci. U.S.A.*, 81(11):3443–3446, Jun 1984.
- [37] Affymetrix. Data Sheet. *GeneChip® Human Mapping 100K Set*, [http://www.affymetrix.com/support/technical/datasheets/100k\\_datasheet.pdf](http://www.affymetrix.com/support/technical/datasheets/100k_datasheet.pdf), 2004-2005.
- [38] Affymetrix. BRLMM. *An Improved Genotype Calling Method for the GeneChip® Human Mapping 500K Array Set.*, [http://www.affymetrix.com/support/technical/whitepapers/brlmm\\_whitepaper.pdf](http://www.affymetrix.com/support/technical/whitepapers/brlmm_whitepaper.pdf), 2006.
- [39] G. R. Abecasis, S. S. Cherny, W. O. Cookson, and L. R. Cardon. Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. *Nat. Genet.*, 30(1):97–101, Jan 2002.
- [40] M. Schuelke. An economic method for the fluorescent labeling of PCR fragments. *Nat. Biotechnol.*, 18(2):233–234, Feb 2000.
- [41] QIAGEN®. QIAGEN® Multiplex PCR Handbook. *QIAGEN®*, 2004.
- [42] Applied Biosystems. Getting Started Guide. *Applied Biosystems 3130/3130xl Genetic Analyzers*, 2004.
- [43] J. Little, M. J. Khoury, L. Bradley, M. Clyne, M. Gwinn, B. Lin, M. L. Lindegren, and P. Yoon. The human genome project is complete. How do we develop a handle for the pump? *Am. J. Epidemiol.*, 157(8):667–673, Apr 2003.
- [44] F. Sanger, S. Nicklen, and A. R. Coulson. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.*, 74(12):5463–5467, Dec 1977.
- [45] M. A. Pujana, J. D. Han, L. M. Starita, K. N. Stevens, M. Tewari, J. S. Ahn, G. Rennert, V. Moreno, T. Kirchhoff, B. Gold, V. Assmann, W. M. Elshamy, J. F. Rual, D. Levine, L. S. Rozek, R. S. Gelman, K. C. Gunsalus, R. A. Greenberg, B. Sobhian, N. Bertin, K. Venkatesan, N. Ayivi-Guedehoussou, X. Sole, P. Hernandez, C. Lazaro,

- K. L. Nathanson, B. L. Weber, M. E. Cusick, D. E. Hill, K. Offit, D. M. Livingston, S. B. Gruber, J. D. Parvin, and M. Vidal. Network modeling links breast cancer susceptibility and centrosome dysfunction. *Nat. Genet.*, 39(11):1338–1349, Nov 2007.
- [46] A. Gonzalez-Neira, J. M. Rosa-Rosa, A. Osorio, E. Gonzalez, M. Southey, O. Sinelnikova, H. Lynch, R. A. Oldenburg, C. J. van Asperen, N. Hoogerbrugge, G. Pita, P. Devilee, D. Goldgar, and J. Benitez. Genomewide high-density SNP linkage analysis of non-BRCA1/2 breast cancer families identifies various candidate regions and has greater power than microsatellite studies. *BMC Genomics*, 8:299, 2007.
- [47] S. Kuznetsov, M. Pellegrini, K. Shuda, O. Fernandez-Capetillo, Y. Liu, B. K. Martin, S. Burkett, E. Southon, D. Pati, L. Tessarollo, S. C. West, P. J. Donovan, A. Nussenzweig, and S. K. Sharan. RAD51C deficiency in mice results in early prophase I arrest in males and sister chromatid separation at metaphase II in females. *J. Cell Biol.*, 176(5):581–592, Feb 2007.
- [48] Y. Liu, M. Tarsounas, P. O’regan, and S. C. West. Role of RAD51C and XRCC3 in genetic recombination and DNA repair. *J. Biol. Chem.*, 282(3):1973–1979, Jan 2007.
- [49] A. Renglin Lindh, N. Schultz, N. Saleh-Gohari, and T. Helleday. RAD51C (RAD51L2) is involved in maintaining centrosome number in mitosis. *Cytogenet. Genome Res.*, 116:38–45, 2007.
- [50] A. Murthy, C. Gonzalez-Agosti, E. Cordero, D. Pinney, C. Candia, F. Solomon, J. Gusella, and V. Ramesh. NHERF, a regulatory cofactor for  $Na^+/H^+$  exchange, is a common interactor for merlin and ERM (MERM) proteins. *J. Biol. Chem.*, 273(3):1273–1276, Jan 1998.
- [51] J. L. Dai, L. Wang, A. A. Sahin, L. D. Broemeling, M. Schutte, and Y. Pan. NHERF ( $Na^+/H^+$  exchanger regulatory factor) gene mutations in human breast cancer. *Oncogene*, 23(53):8681–8687, Nov 2004.
- [52] Y. Pan, E. J. Weinman, and J. L. Dai.  $Na^+/H^+$  exchanger regulatory factor 1 inhibits platelet-derived growth factor signaling in breast cancer cells. *Breast Cancer Res.*, 10(1):R5, 2008.
- [53] Y. Pan, L. Wang, and J. L. Dai. Suppression of breast cancer cell growth by  $Na^+/H^+$  exchanger regulatory factor 1 (NHERF1). *Breast Cancer Res.*, 8(6):R63, 2006.
- [54] Y. Ding, L. Wang, L. K. Su, J. A. Frey, R. Shao, K. K. Hunt, and D. H. Yan. Antitumor activity of IFIX, a novel interferon-inducible HIN-200 gene, in breast cancer. *Oncogene*, 23(26):4556–4566, Jun 2004.
- [55] Y. Ding, J. F. Lee, H. Lu, M. H. Lee, and D. H. Yan. Interferon-inducible protein IFIXalpha1 functions as a negative regulator of HDM2. *Mol. Cell. Biol.*, 26(5):1979–1996, Mar 2006.
- [56] E. I. Campos, M. Y. Chin, W. H. Kuo, and G. Li. Biological functions of the ING family tumor suppressors. *Cell. Mol. Life Sci.*, 61:2597–2613, Oct 2004.
- [57] G. S. Nouman, J. J. Anderson, J. Lunec, and B. Angus. The role of the tumour suppressor p33 ING1b in human neoplasia. *J. Clin. Pathol.*, 56(7):491–496, Jul 2003.

- [58] T. Toyama, H. Iwase, P. Watson, H. Muzik, E. Saettler, A. Magliocco, L. DiFrancesco, P. Forsyth, I. Garkavtsev, S. Kobayashi, and K. Riabowol. Suppression of ING1 expression in sporadic breast cancer. *Oncogene*, 18(37):5187–5193, Sep 1999.
- [59] F. Kerangueven, L. Essioux, A. Dib, T. Noguchi, F. Allione, J. Geneix, M. Longy, R. Lidereau, F. Eisinger, and M. J. Pebusque. Loss of heterozygosity and linkage analysis in breast carcinoma: indication for a putative third susceptibility gene on the short arm of chromosome 8. *Oncogene*, 10(5):1023–1026, Mar 1995.
- [60] S. Seitz, K. Rohde, E. Bender, A. Nothnagel, K. Kolble, P. M. Schlag, and S. Scherneck. Strong indication for a breast cancer susceptibility gene on chromosome 8p12-p22: linkage analysis in German breast cancer families. *Oncogene*, 14(6):741–743, Feb 1997.
- [61] T. Kainu, S. H. Juo, R. Desper, A. A. Schaffer, E. Gillanders, E. Rozenblum, D. Freas-Lutz, D. Weaver, D. Stephan, J. Bailey-Wilson, O. P. Kallioniemi, M. Tirkkonen, K. Syrjakoski, T. Kuukasjarvi, P. Koivisto, R. Karhu, K. Holli, A. Arason, G. Johannesdottir, J. T. Bergthorsson, H. Johannsdottir, V. Egilsson, R. B. Barkardottir, O. Johannsson, K. Haraldsson, T. Sandberg, E. Holmberg, H. Gronberg, H. Olsson, A. Borg, P. Vehmanen, H. Eerola, P. Heikkila, S. Pyrhonen, and H. Nevanlinna. Somatic deletions in hereditary breast cancers implicate 13q21 as a putative novel breast cancer susceptibility locus. *Proc. Natl. Acad. Sci. U.S.A.*, 97(17):9603–9608, Aug 2000.
- [62] N. Rahman, M. D. Teare, S. Seal, H. Renard, J. Mangion, C. Cour, D. Thompson, Y. Shugart, D. Eccles, P. Devilee, H. Meijers, K. L. Nathanson, S. L. Neuhausen, B. Weber, J. Chang-Claude, D. F. Easton, D. Goldgar, and M. R. Stratton. Absence of evidence for a familial breast cancer susceptibility gene at chromosome 8p12-p22. *Oncogene*, 19(36):4170–4173, Aug 2000.
- [63] D. Thompson, C. I. Szabo, J. Mangion, R. A. Oldenburg, F. Odefrey, S. Seal, R. Barfoot, K. Kroeze-Jansema, D. Teare, N. Rahman, H. Renard, G. Mann, J. L. Hopper, S. S. Buys, I. L. Andrulis, R. Senie, M. B. Daly, D. West, E. A. Ostrander, K. Offit, T. Peretz, A. Osorio, J. Benitez, K. L. Nathanson, O. M. Sinilnikova, E. Olah, Y. J. Bignon, P. Ruiz, M. D. Badzioch, H. F. Vasen, A. P. Futreal, C. M. Phelan, S. A. Narod, H. T. Lynch, B. A. Ponder, R. A. Eeles, H. Meijers-Heijboer, D. Stoppa-Lyonnet, F. J. Couch, D. M. Eccles, D. G. Evans, J. Chang-Claude, G. Lenoir, B. L. Weber, P. Devilee, D. F. Easton, D. E. Goldgar, and M. R. Stratton. Evaluation of linkage of breast cancer to the putative BRCA3 locus on chromosome 13q21 in 128 multiple case families from the Breast Cancer Linkage Consortium. *Proc. Natl. Acad. Sci. U.S.A.*, 99(2):827–831, Jan 2002.
- [64] P. Huusko, S. H. Juo, E. Gillanders, L. Sarantaus, T. Kainu, P. Vahteristo, M. Allinen, M. Jones, K. Rapakko, H. Eerola, C. Markey, P. Vehmanen, D. Gildea, D. Freas-Lutz, C. Blomqvist, J. Leisti, G. Blanco, U. Puistola, J. Trent, J. Bailey-Wilson, R. Winqvist, H. Nevanlinna, and O. P. Kallioniemi. Genome-wide scanning for linkage in Finnish breast cancer families. *Eur. J. Hum. Genet.*, 12(2):98–104, Feb 2004.
- [65] A. Bergman, P. Karlsson, J. Berggren, T. Martinsson, K. Bjorck, S. Nilsson, J. Wahlstrom, A. Wallgren, and M. Nordling. Genome-wide linkage scan for breast cancer susceptibility loci in Swedish hereditary non-BRCA1/2 families: suggestive linkage to 10q23.32-q25.3. *Genes Chromosomes Cancer*, 46(3):302–309, Mar 2007.

- 
- [66] P. Smith, L. McGuffog, D. F. Easton, G. J. Mann, G. M. Pupo, B. Newman, G. Chenevix-Trench, C. Szabo, M. Southey, H. Renard, F. Odefrey, H. Lynch, D. Stoppa-Lyonnet, F. Couch, J. L. Hopper, G. G. Giles, M. R. McCredie, S. Buys, I. Andrulis, R. Senie, D. E. Goldgar, R. Oldenburg, K. Kroeze-Jansema, J. Kraan, H. Meijers-Heijboer, J. G. Klijn, C. van Asperen, I. van Leeuwen, H. F. Vasen, C. J. Cornelisse, P. Devilee, L. Baskcomb, S. Seal, R. Barfoot, J. Mangion, A. Hall, S. Edkins, E. Rapley, R. Wooster, J. Chang-Claude, D. Eccles, D. G. Evans, P. A. Futreal, K. L. Nathanson, B. L. Weber, N. Rahman, and M. R. Stratton. A genome wide linkage search for breast cancer susceptibility genes. *Genes Chromosomes Cancer*, 45(7):646–655, Jul 2006.

## 7 ABBILDUNGSVERZEICHNIS

1	Brustkrebs susceptibility. . . . .	11
2	Logik des positionellen Kandidatengenverfahrens. . . . .	14
3	Ausschnitt aus dem Stammbaum der untersuchten Familie. . . . .	18
4	Amplifikationsschema der PCR-Methode nach Schuelke (2000) [40]. . . . .	27
5	Heterozygote Familiensituation. . . . .	31
6	Homozygote Familiensituation. . . . .	31
7	Stotterbanden eines heterozygoten Dinukleotidrepeats. . . . .	32
8	Ablauf der Sequenzierung. . . . .	33
9	Haplotypenanalyse des Chromosom X. . . . .	39
10	Haplotypenanalyse des Chromosom 17. . . . .	40
11	Parametrische Analyse III-1 "unknown". . . . .	41
12	Parametrische Analyse III-5 "unknown". . . . .	41
13	Parametrische Analyse III-6 "unknown". . . . .	42
14	Parametrische Analyse III-7 "unknown". . . . .	42
15	Parametrische Analyse III-9 "unknown". . . . .	42
16	Parametrische Analyse III-10 "unknown". . . . .	43
17	Haplotypenanalyse 1q Individuum III-1 "unknown". . . . .	44
18	Haplotypenanalyse 17q Individuum III-1 "unknown". . . . .	45
19	Haplotypenanalyse 20q Individuum III-1 "unknown". . . . .	46
20	Haplotypenanalyse 7q Individuum III-5 "unknown". . . . .	47
21	Haplotypenanalyse 8p Individuum III-6 "unknown". . . . .	48
22	Haplotypenanalyse 13q Individuum III-6 "unknown". . . . .	49
23	Haplotypenanalyse 19 Individuum III-7 "unknown". . . . .	50

## 8 ANHANG

### 8.1 Abkürzungen

A	Adenin
AT	Ataxia teleangiectasia
ATM	Ataxia teleangiectasia mutated protein
BACH1	BRCA1-interagierendes Protein C-terminale Helicase 1
bp	Basenpaar(e)
BRCA1	Breast cancer Typ 1 Suszeptibilitätsprotein
BRCA2	Breast cancer Typ 2 Suszeptibilitätsprotein
BRIP1	BRCA1-interagierendes Protein 1
bzw.	Beziehungsweise
C	Cytosin
°C	Grad Celsius
CDH1	Cadherin 1, E-cadherin (epithelial)
CD1C	CD1C Antigen
CGH	Comparative genomic hybridization
CHEK2	Checkpoint kinase 2
CS	Cowden Syndrom
CSP8	Caspase 8
cM	Centi-Morgan
dATP	Desoxyadenosintriphosphat
dCTP	Desoxycytosintriphosphat
ddNTP	Didesoxynukleotid
ddATP	Didesoxyadenosintriphosphat
ddCTP	Didesoxycytosintriphosphat
ddGTP	Didesoxyguanidintriphosphat
ddTTP	Didesoxythymidintriphosphat
DDX39	DEAD box protein 39
dGTP	Desoxyguanidintriphosphat
d.h.	Das heißt
DHPLC	Denaturing High Performance Liquid Chromatography
DNA	Desoxyribonukleinsäure
dNTP	Desoxynukleotid
dTTP	Desoxythymidintriphosphat
EDTA	Ethylendiamintetraessigsäure
EST	Expressed sequenced tag
FA	Fanconi Anämie
FGFR2	Fibroblast growth factor receptor 2
g	Gramm
G	Guanin
H <sub>0</sub>	Nullhypothese
H <sub>1</sub>	Hypothese 1
H <sub>2</sub> O	Wasser
HCl	Chlorwasserstoff
HDGCS	Hereditäres diffuses Gastric Cancer Syndrom
IBD	Identical by descent
IBS	Identical by state



---

IFIX	Interferon-inducible Protein X
ILF3	Interleukin enhancer-binding factor 3
ING1	Inhibitor of growth Protein 1
kb	Kilobase
KCl	Kaliumchlorid
L	Likelihood
l	Liter
LFS	Li-Fraumeni Syndrom
LKB1	Serine/threonine-protein kinase 11
LOD	Logarithm of the odds
LOH	Loss of heterozygosity
LSP1	Lymphocyte-specific protein 1
M	Molar (mol/l)
MAP3K1	Mitogen-activated protein kinase kinase kinase 1
Mb	Megabase
MgCl <sub>2</sub>	Magnesiumchlorid
MgSO <sub>4</sub>	Magnesiumsulfat
min	Minute(n)
ml	Milliliter
MLH1	MutL protein homolog 1
mLOD	multipoint LOD-Wert
mM	Millimolar
mögl.	Mögliche/möglicher/mögliches
MSH2	MutS Proteinhomolog 2
MYBL2	V-myb avian myeloblastosis viral oncogene homolog-like 2
µl	Mikroliter
NaAc	Natriumacetat
ng	Nanogramm
NHERF1	$Na^+/H^+$ exchange regulatory cofactor
nM	Nanomolar
NOL	Nicht-parametrischer LOD-Wert
NOLOD	Nicht-parametrischer LOD-Wert
p	Kurzer Arm des Chromosoms
PALB2	Partner and localizer of BRCA2
PAXIP1L	PAX-interacting protein 1
PCR	Polymerase chain reaction
PJS	Peutz-Jeghers Syndrom
PYHIN1	Pyrin and HIN domain family member 1
pMol	Pikomol
PTEN	Phosphatase and tensin homolog
p53	Tumorprotein 53
q	Langer Arm des Chromosoms
R	Likelihood ratio, odds ratio
RAD51	DNA-Reparaturprotein RAD51 homolog 1
RAD51C	RAD51 homolog C
RAD51L2	RAD51-like protein 2
Rb1	Retinoblastoma Suszeptibilitätsprotein 1
rpm	Rounds per minute
sec	Sekunde(n)

---

SLC9A3R1	Solute carrier family 9 (sodium/hydrogen exchanger), Isoform A3 regulatory factor 1
SNP	Single nucleotide polymorphism
SSR	Simple sequence repeat
STK11	Serine/threonine-protein kinase 11
STR	Short tandem repeat
T	Thymin
T <sub>A</sub>	Annealingtemperatur
T <sub>m</sub>	Schmelztemperatur
TFDP1	Transcription factor Dp-1
TOX3	TOX high mobility group box family member 3
TP53	Tumorprotein 53
Tris	Tris(hydroxymethyl)-aminomethan
$\theta$ (theta)	Rekombinationsfraktion
v.a.	Vor allem
U	Unit, Maßeinheit der Enzymaktivität
UBE2C	Ubiquitin-conjugating enzyme E2 C
UK	United Kingdom
Z( $\theta$ )	LOD-Wert in Abhängigkeit von $\theta$
Z <sub>mean</sub>	Nicht-parametrische LOD-Wert in Merlin Outputdatei
ZNF43	Zinkfinger Protein 43
$\omega$	Abstand auf genetischer Karte

## 8.2 Auswertungen

Tabelle 1: Mikrosatelliten-Marker

Tabelle 2: Sequenzierte Exone

Tabelle 3: Gene in der Kandidatenregion Xq27.2-q28 (140384457-154913754)

Tabelle 4: Gene in der Kandidatenregion 17q21.33-q25.1 (46654293-70734460)

Tabelle 5: Gene in der Kandidatenregion 1q23.1-q23.3 (156199127-160592447)

Tabelle 6: Gene in der Kandidatenregion 20q12-q13.2 (40749905-49190813)

Tabelle 7: Gene in der Kandidatenregion 7q36.2-q36.3 (153661756-156156741)

Tabelle 8: Gene in der Kandidatenregion 8p23.3 (1-2117883)

Tabelle 9: Gene in der Kandidatenregion 13q32.1-q34 (96484285-114142980)

Tabelle 10: Gene in der Kandidatenregion 19p13.2-q12 (10556621-35934942)

Tabelle 1: Mikrosatelliten-Marker

Chr.	Markenname	Sequenz der Primer M13-Forward Primer: M13-F Reverse Primer: R	Hetero- zygotität	Produkt- größe	Repeat	T <sub>A</sub>	PCR Bed.	Chr. Start (bp)	Chr. End (bp)
1	GATA43A04 (D1S1653)	M13-F: TGTAAACGACGGCCAGTCAGCAAGATCACCCTCAT F: GTCAGCAAGATCACCCTCAT R: CTGGAAAGCCTGTAGGAAGA	0,68	381 bp	Tetra	58°C	-	156199127	156199507
1	NIB1152 (D1S1167)	M13-F: TGTAAACGACGGCCAGTAAGCATAGCATGGCCAAAT F: AAGCATAGCATGGCCAAAT R: TTCTTTTCAGTGTGTAGGACAA	0,84	145 bp	Di	58°C	-	158275726	158275870
1	AFMA339WH1 (D1S2707)	M13-F: TGTAAACGACGGCCAGTCACACACACATTCGGCTTT F: CACACACACATTCGGCTTT R: GGAAAGAGGAGGGTTTCAGAG	0,82	149 bp	Di	61°C	+	158339025	158339173
7	AFMA053XB5 (D7S2546)	M13F: TGTAAACGACGGCCAGTGCCAGGTAATATTGGAGCAA F: GCCAGGTAATATTGGAGCAA R: GCCAGGTCATCTCTTTTCCT	0,76	273 bp	Di	56°C	+	153838753	153839025
7	MFD442- GTTT002	M13-F: TGTAAACGACGGCCAGTATCTCTCAGCAGCGGTCAAAAT F: ATCTCTCAGCAGCGTCAAAAT R: ATGACCCAAAACCTCAGACCA	0,75	232 bp	Tetra	58°C	-	155681848	155682079
8	21 × TG	M13-F: TGTAAACGACGGCCAGTGCCCTTTTCAATGTCGTCTCT F: TCCCGTTTCAATGTCGTCTCT R: AGAGCATCGAAGCAAACT	unknown	142 bp	Di	61°C	-	961185	961326
8	AFM143XD8 (D8S264)	M13-F: TGTAAACGACGGCCAGTGGGAGATTTCGGAACATCTG F: GGGAGATTTCGGAACATCTG R: ACCAAGACCTGAGTCAGCAT	0,83	144 bp	Di	56°C	-	2117740	2117883
13	AFM338WD9 (D13S1298)	M13F: TGTAAACGACGGCCAGTTCTTCCCTTTGACTTTCCA F: TTCTTCCCTTTGACTTTCCA R: CAGTGGGACTTCTCTGACC	0,84	165 bp	Di	58°C	-	97823754	97823918
13	AFM155YB6 (D13S159)	M13-F: TGTAAACGACGGCCAGTCTCCAGTTGTTGGGATTTA F: CTCCTCCAGTTGTTGGGATTTA R: TACCCAAAACAGCTAGCACCA	0,88	197 bp	Di	61°C	+	97851570	97851766
13	ATA261D07 (D13S779)	M13-F: TGTAAACGACGGCCAGTAGTCACGTCCACATTCACAA F: AGTCACGTCCACATTCACAA R: CTGTATGCAGTCAGCTCTGG	0,7	462 bp	Tri	58°C	-	100301667	100302128
17	AFM338XG5 (D17S1606)	M13-F: TGTAAACGACGGCCAGTGCCCTGTGTGTTCAATGTTGAT F: GCCTGTGTGTTCAATGTTGAT R: AAGGGGCACAGTTAAGGTAA	0,88	198 bp	Di	58°C	+	52958668	52958865
17	GATA49C09 (D17S1290)	M13-F: TGTAAACGACGGCCAGTGCACCTCACACCAAGAATTT F: TGCACCTCACCAAGAATTT R: TAAAGGACTTCTCCCTGTGC	0,84	445 bp	Tetra	56°C	-	53686148	53686592
19	UT705 (D19S394)	M13-F: TGTAAACGACGGCCAGTGAAGAGCCAGGTGCAAT F: GAAGAGGCCAGGTGCAAT R: CCTGACTACACAGGCACAATC	0,86	445 bp	Tetra	60°C	-	10556621	10557065
19	ATA37G08	M13-F: TGTAAACGACGGCCAGTGGGTTTCAGGGAAATTTATG F: TGGGTTTCAGGAAATTTATG R: CAGCATGAAAGAAAAGATGC	0,71	139 bp	Tri	56°C	Q-Sol	15953818	15953956
19	GGA21A04! (D19S1037)	M13-F: TGTAAACGACGGCCAGTGAACAAAGGTCCTGCTCTGA F: GGAACAAAGGTCCTGCTCTGA R: AGGGAGGCTGTCTCTGTGT	0,78	234 bp	Tetra	58°C	+	17937712	17937945

Chr.	Markername	Sequenz der Primer M13-Forward Primer: M13-F Reverse Primer: R	Heterozygosität	Produktgröße	Repeat	T <sub>A</sub>	PCR Bed.	Chr. Start (bp)	Chr. End (bp)
20	AFMB298YH5 (D20S887)	M13-F: TGTAACACGACGGCCAGTTGAGTGGGTTTCAGACTCCTT F: TGAGTGGGTTTCAGACTCCTT R: ATGGCCTACGTCGATTAAAC	0,82	308 bp	Di	58°C	-	47106659	47106966
20	AAT269	M13-F: TGTAACACGACGGCCAGTTACCCCTCAGAACTGACTCTGG F: ACCCTCAGAACTGACTCTGG R: GGCTGGATTCTGTATCCACT	0,75	229 bp	Tri	61°C	-	47389917	47390145
X	AFMB018WD9 (DXS8043)	M13-F: TGTAACACGACGGCCAGTTGCAAAAGAGTACAGGCAGGAC F: GCAAAAGAGTACAGGCAGGAC R: CTCAGAAACATTTGGTTAGGC	0,9	173 bp	Di	58°C	-	143836222	143836394
X	TATC043	M13-F: TGTAACACGACGGCCAGTACAGTGGTTGCTCTCAAT F: GTCACGTGGTTGCTCTCAAT R: AGCTTGGGCTAGTCAATCAA	0,75	145 bp	Tetra	58°C	+	145513535	145513679
X	AFMA065WD9 (DXS8103)	M13-F: TGTAACACGACGGCCAGTTCTTTCCCTTCCCTCCAGTT F: TTCTTCCCTTCCCTCCAGTT R: TGGCCAAATTCCTGGTCAC	0,88	151 bp	Di	58°C	-	149865767	149865917

Tabelle 2: Sequenzierte Exone

Exon	Individuen	Sequenz Forward Primer	Sequenz Reverse Primer	Länge (bp)	PCR- Bed.	T <sub>a</sub>
RAD51C-Ex1	III-5, II-5, II-6	CGGAATGGTGCATAAAGTGTG	AGGCGAGAGAACGAAAGACTG	393	-	58°C
RAD51C-Ex2	III-5, II-5, II-6	CTCCACTCTGAGTACACTGTT	TGGTTTCTGACGATAGTACAAA	451	-	58°C
RAD51C-Ex3	III-5, II-5, II-6	TCTGTTGCCTTGGGGAGTAT	TCCAAATGGAGTGTGCTGA	389	-	58°C
RAD51C-Ex4	III-5, II-5, II-6	AAGAATTTTGTATTCACAAAGGAG	GGCTTTGACTTTGATTTTATGCG	350	-	58°C
RAD51C-Ex5	III-5, II-5, II-6	GTAAAGGTTTGGATTAAAGAAGAGG	TGTCAGGCAACCGGTATTTT	500	-	58°C
RAD51C-Ex6	III-5, II-5, II-6	ggcagactggctctACTTTGAT	tggtctggccACTCAATAAA	336	-	58°C
RAD51C-Ex7	III-5, II-5, II-6	GATTAATTTTATCATAGAGGCGTTC	GGTGATATCAGACAAGGCAAGA	289	-	58°C
RAD51C-Ex8	III-5, II-5, II-6	ACGGGTAAATTTGAAGGGTGT	CCTCAAATATGCTTGTGCGC	348	-	58°C
RAD51C-Ex9.2	III-5, II-5, II-6	gcttgccctTAGAATAAAGT	CCACAATGAGATCAGCTTTCC	488	-	60°C
PYHIN1-Ex1	III-5, II-5, II-6	TCTCTAACAGGATGGAAGCAA	CAC TGCCCTGAAAGACCAAT	407	+	55°C
PYHIN1-Ex2	III-5, II-5, II-6	tcacatagGCATACATCTTCTTT	CATGCCAGTGAAGTCAACA	480	+	55°C
PYHIN1-Ex3	III-5, II-5, II-6	CCCTCAAGCAGGAGCTAAGA	CAGATTGAGTGGAGAGCCA	317	-	60°C
PYHIN1-Ex4	III-5, II-5, II-6	TTGGCCTGGGATGTAGTTA	TTTAGTGGCTGATGATTTGATGC	403	-	60°C
PYHIN1-Ex5	III-5, II-5, II-6	GTTGTGCTTTGAGGTCACTG	CCATTTAGCAGATCAGGGCT	581	-	60°C
PYHIN1-Ex6	III-5, II-5, II-6	GGAGTTGCTGTCTTGCATCTT	TAAGCGCAAGGGCATTTATTT	432	-	60°C
PYHIN1-Ex7	III-5, II-5, II-6	GGCAATCTTTGAAGGTAATGACA	GAGATTGTACAGATTAGGATTAGA	397	-	60°C
PYHIN1-Ex8	III-5, II-5, II-6, III-9, II-12, II-13	CCAGAGAGATTAGCAGAGGGA	AAGGACTTAGAATCCAGTGGCA	482	-	60°C
PYHIN1-Ex9	III-5, II-5, II-6	GGCAGTGTGATTGAAGAAAGA	TTGCAGGCCACAGTTTATTT	500	-	60°C
ING1a-Ex1	III-9, II-12, II-13	CGAGGGCTTTGCATTTTG	AGGACCTTGGCGAATCAA	410	AccuPrime*	58,5°C
ING1c-Ex1	III-9, II-12, II-13	GAGCTTTGGAGGGTAGAGG	AAACCAAGTCTCGGGG	491	Q-Sol.	62°C
ING1d-Ex1_1	III-9, II-12, II-13	GGAAAGTACTAGAGCTCTGCG	CGGGATCACTGCTACTGCTA	574	Q-Sol.	62°C
ING1d-Ex1_2	III-9, II-12, II-13	GCCTTGGATTTGGTCTCTCTC	CCATGAACCTTTGCTGTCTGG	592	+	62°C
ING1-Ex2_1	III-9, II-12, II-13	GCTGTGGGCTTTGTTCT	CCTTGGAGCGCTTCTTCT	575	Q-Sol.	58,5°C
ING1-Ex2_2	III-9, II-12, II-13	ACAAGCCCAACAGCAAG	CACCTCCTTGCACTCAACAA	497	-	60°C
SLC9A3R1-Ex1_1	III-9, II-12, II-13	GTCTCTGGGACACCTGCTT	GCGAGGCTCATTTTCGTT	564	AccuPrime*	60°C
SLC9A3R1-Ex1_2	III-9, II-12, II-13	AAAACGTGGAGAGGAGAGC	CGCAGCTCGAGTTTCAGA	366	AccuPrime*	60°C
SLC9A3R1-Ex2	III-9, II-12, II-13	TGCTGTGAGGGATCTAGGC	CAGCTCCAGGTTGGCTATG	419	-	60°C
SLC9A3R1-Ex3	III-9, II-12, II-13	TCAGTATCCCCAGGTGTGTA	CAGAAACCCACTGCCACTC	489	AccuPrime*	60°C
SLC9A3R1-Ex4	III-9, II-12, II-13	ATAATCTGCCCAACCCCAAC	GAAGAAAGAGGTGGAACCTGG	264	-	62°C
SLC9A3R1-Ex5	III-9, II-12, II-13	TAGAACCAGCTCGCTCT	AGGGAAGCCAGGTCAGAAC	428	AccuPrime*	60°C
SLC9A3R1-Ex6	III-9, II-12, II-13	TGAGCCCGATTTCTGTCTT	TTTGTCAAGAAAGGGGATGT	446	-	58°C

AccuPrime\*: AccuPrime<sup>TM</sup> GC-Rich DNA Polymerase

Tabelle 3: Gene in der Kandidatenregion Xq27.2-q28 (140384457-154913754)

Kandidaten region	GenSymbol	RefSeq- Nummer	Transkription Start	Transkription Ende	Exone	Name (englisch)	kurze Funktionsbeschreibung	Rang
Xq	ABCD1	NM_000033	152643529	152663374	10	ATP-binding cassette, sub-family D (ALD)	Peroxisomales Membranprotein, Transport und Katabolismus von Fettsäuren	3
Xq	AFF2	NM_002025	147389830	147889899	21	fragile X mental retardation 2	unbekannt	4
Xq	ARD1A	NM_003491	152848570	152853662	8	N-acetyltransferase, homolog of S. cerevisiae	Acetyltransferase	3
Xq	ARDHAP4	NM_001666	152826024	152844892	22	Rho GTPase activating protein 4	Inhibitorischer Effekt auf Stressfaserorganisation	3
Xq	ATP2B3	NM_001001344	152454773	152501581	20	plasma membrane calcium ATPase 3 isoform 3b	Calcium-ATPase	3
Xq	ATP2B3	NM_021949	152454773	152501581	21	plasma membrane calcium ATPase 3 isoform 3a	Calcium-ATPase	3
Xq	ATP6AP1	NM_001183	153310171	153318056	10	ATPase, H+ transporting, lysosomal accessory	Vakuolar-ATPase	3
Xq	AVPR2	NM_000054	152823563	152825834	3	arginine vasopressin receptor 2	Rezeptor für Arginin/Vasopressin	3
Xq	BCAP31	NM_005745	152619145	152643081	8	B-cell receptor-associated protein 31	Rolle im intrazellulären Proteintransport und CASP8-vermittelter Apoptose	3
Xq	BGN	NM_001711	152413604	152428198	8	biglycan preproprotein	Protein des Bindegewebemetabolismus	3
Xq	BRCC3	NM_001018055	153952903	154004543	11	BRCA1/BRCA2-containing complex, subunit 3	Untereinheit des BRCA1/BRCA2-enthaltenden Komplexes (BRCC), Rolle in zellulärer Antwort auf Strahlung und im G2/M-Checkpoint	2
Xq	BRCC3	NM_024332	153952903	154004543	12	BRCA1/BRCA2-containing complex, subunit 3	Untereinheit des BRCA1/BRCA2-enthaltenden Komplexes (BRCC), Rolle in zellulärer Antwort auf Strahlung und im G2/M-Checkpoint	2
Xq	CD99L2	NM_031462	149685466	149817837	11	CD99 antigen-like 2 isoform E3'E4'E3-E4	unbekannt	4
Xq	CD99L2	NM_134445	149685466	149817837	8	CD99 antigen-like 2 isoform E4	unbekannt	4
Xq	CD99L2	NM_134446	149685466	149817837	9	CD99 antigen-like 2 isoform E3-E4	unbekannt	4
Xq	CETN2	NM_004344	151746526	151749957	5	caltractin	Rolle in Struktur und Funktion des Mikrotubulus-Organisationszentrums	3
Xq	CLIC2	NM_001289	154158693	154217180	6	chloride intracellular channel 2	Chlorid-Kanal	3
Xq	CINGA2	NM_005140	150653873	150664692	7	cyclic nucleotide gated channel alpha 2	Protein der olfaktorischen Signaltransduktion	3
Xq	CSAG1	NM_153478	151653883	151660174	5	chondrosarcoma associated gene 1 isoform a	Tumorantigen, Funktion unbekannt	4
Xq	CSAG1	NM_153479	151653883	151660174	4	chondrosarcoma associated gene 1 isoform b	Tumorantigen, Funktion unbekannt	4
Xq	CSAG2	NM_00180848	151627398	151628403	2	hypothetical protein LOC728461	Tumorantigen, Funktion unbekannt	4
Xq	CSAG2	NM_004909	151678389	151679394	2	CSAG family, member 2	Tumorantigen, Funktion unbekannt	4
Xq	CSAG3A	NM_203311	151627398	151633694	5	CSAG family, member 3A	Tumorantigen, Funktion unbekannt	4
Xq	CSAG3A	NM_203311	151627398	151633694	5	CSAG family, member 3A	Tumorantigen, Funktion unbekannt	4
Xq	CTAG1A	NM_139250	153499058	153500716	3	cancer/testis antigen 1A	Tumorantigen, Funktion unbekannt	4
Xq	CTAG1B	NM_001327	153466611	153468269	3	cancer/testis antigen 1B	Tumorantigen, Funktion unbekannt	4
Xq	CTAG2	NM_020994	153533444	153535036	2	cancer/testis antigen 2 isoform LAGE-1b	Tumorantigen, Funktion unbekannt	4
Xq	CTAG2	NM_172377	153533444	153535036	3	cancer/testis antigen 2 isoform LAGE-1a	Tumorantigen, Funktion unbekannt	4
Xq	CXorf1	NM_004709	144716619	144719062	1	hypothetical protein LOC9142	unbekannt	4
Xq	CXorf40A	NM_178124	148430442	148436760	5	chromosome X open reading frame 40	mögl. Rolle im Zellschutz während inflammatorischer Reaktion	4
Xq	CXorf40B	NM_001013845	148851072	148857374	5	hypothetical protein LOC541578	unbekannt	4
Xq	CXorf6	NM_005491	149364377	149433104	7	hypothetical protein LOC10046	mögl. Rolle in Gonadenfunktion	4
Xq	DKC1	NM_001363	153644343	153659154	15	dykerin	Rolle in Biogenese der Ribosomen und Erhaltung der Telomere	3
Xq	DNASE1L1	NM_001009932	153282772	153293621	10	deoxyribonuclease I-like 1 precursor	unbekannt	4
Xq	DNASE1L1	NM_001009933	153282772	153293621	9	deoxyribonuclease I-like 1 precursor	unbekannt	4
Xq	DNASE1L1	NM_001009934	153282772	153293621	8	deoxyribonuclease I-like 1 precursor	unbekannt	4
Xq	DNASE1L1	NM_006730	153282772	153290769	9	deoxyribonuclease I-like 1 precursor	unbekannt	4

Xq	DUSP9	NM_001395	152561181	152569971	4	dual specificity phosphatase 9	Negative Regulation von Proteinen der MAP-Kinasen Superfamilie	3
Xq	EMD	NM_000117	153260980	153263075	6	emerin	Nukleäres Membranprotein, Verankerung der Membran im Zytoskelett	3
Xq	F8	NM_000132	153717257	153904192	26	coagulation factor VIII isoform a precursor	Faktor VIII der Koagulation	3
Xq	F8	NM_019863	153717257	153767771	5	coagulation factor VIII isoform b precursor	Faktor VIII der Koagulation	3
Xq	F8A3	NM_001007524	153767828	153769330	1	coagulation factor VIII-associated (intronic)	unbekannt	4
Xq	F8A3	NM_001007524	154264942	154266644	1	coagulation factor VIII-associated (intronic)	unbekannt	4
Xq	F8A3	NM_001007524	154339768	154341470	1	coagulation factor VIII-associated (intronic)	unbekannt	4
Xq	FAM3A	NM_021806	153387699	153397567	9	family 3, member A protein	unbekannt	4
Xq	FAM50A	NM_004699	153325701	153332185	13	XAP-5 protein	unbekannt	4
Xq	FAM58A	NM_152274	152506576	152517775	5	hypothetical protein LOC92002	unbekannt	4
Xq	FATE1	NM_033085	150835163	150842320	5	fetal and adult testis expressed transcript	unbekannt	4
Xq	FLJ00038	NM_182905	154904539	154908524	9	hypothetical protein LOC375690	unbekannt	4
Xq	FLNA	NM_001456	153230090	153256123	48	filamin A, alpha	Regulation der Organisierung des Aktin-Zytoskeletts	3
Xq	FMR1	NM_002024	146801200	146840303	18	fragile X mental retardation 1	RNA-bindendes Protein, mRNA-Transport	3
Xq	FMRINB	NM_152578	146870540	146915876	6	fragile X mental retardation 1 neighbor	Tumorantigen, Funktion unbekannt	4
Xq	FUNDQ2	NM_023934	153908257	153938385	5	FUN14 domain containing 2	unbekannt	4
Xq	G6PD	NM_000402	153412799	153428427	13	glucose-6-phosphate dehydrogenase isoform a	Glukose-6-Phosphat-Dehydrogenase	3
Xq	G6PD	NM_001042351	153412799	153428981	13	glucose-6-phosphate dehydrogenase isoform b	Glukose-6-Phosphat-Dehydrogenase	3
Xq	GAB3	NM_001081573	153556720	153632542	10	Gab3 protein isoform 1	Rolle in Makrophagendifferenzierung	3
Xq	GAB3	NM_080612	153556720	153632542	10	Gab3 protein isoform 2	Rolle in Makrophagendifferenzierung	3
Xq	GABRA3	NM_000808	151087185	151370486	11	gamma-aminobutyric acid A receptor, alpha 3	GABA-Rezeptor	3
Xq	GABRE	NM_004961	150872251	150893807	9	gamma-aminobutyric acid (GABA) A receptor	GABA-Rezeptor	3
Xq	GABRE	NM_021984	150872251	150893807	10	gamma-aminobutyric acid (GABA) A receptor	GABA-Rezeptor	3
Xq	GABRE	NM_021987	150872251	150893807	11	gamma-aminobutyric acid (GABA) A receptor	GABA-Rezeptor	3
Xq	GABRE	NM_021990	150872251	150893807	11	gamma-aminobutyric acid (GABA) A receptor	GABA-Rezeptor	3
Xq	GABRQ	NM_018558	151557292	151572481	9	gamma-aminobutyric acid (GABA) receptor, theta	GABA-Rezeptor	3
Xq	GDI1	NM_001493	153318714	153325009	11	GDP dissociation inhibitor 1	Regulation der GDP-GTP-Austauschreaktion in Proteinen der Rab-Familie	3
Xq	GPR50	NM_004224	150095782	150100595	3	G protein-coupled receptor 50	unbekannt	4
Xq	H2AFB2	NM_001017991	154263621	154264138	1	H2A histone family, member B2	Nukleosomkomponente Histon H2A	3
Xq	H2AFB2	NM_001017991	154342273	154342790	1	H2A histone family, member B2	Nukleosomkomponente Histon H2A	3
Xq	H2AFB3	NM_080720	153766510	153767027	1	H2A histone family, member B3	Nukleosomkomponente Histon H2A	3
Xq	HCFP1	NM_005334	152866201	152890013	26	host cell factor 1	Involviert in Reaktion auf Herpes simplex Virusinfektion	3
Xq	HMGCB3	NM_005342	149902420	149909906	5	high-mobility group box 3	unbekannt	4
Xq	HSPX1	NM_016153	148481977	148663581	3	heat shock transcription factor family, X linked	Transkriptionsfaktor	4
Xq	HSPX1	NM_016153	148484725	148666329	3	heat shock transcription factor family, X linked	Transkriptionsfaktor	4
Xq	IDH3G	NM_004135	152704414	152713161	13	isocitrate dehydrogenase 3 (NAD+) gamma isoform	Isocitrat-Dehydrogenase	3
Xq	IDH3G	NM_174869	152704414	152713161	12	isocitrate dehydrogenase 3 (NAD+) gamma isoform	Isocitrat-Dehydrogenase	3
Xq	IDS	NM_000202	148368205	148394698	9	iduronate-2-sulfatase isoform a precursor	Iduronat-2-Sulfatase	3



Xq	IDS	NM_006123	148376568	148394698	8	iduronate-2-sulfatase isoform b precursor	Iduronat-2-Sulfatase	3
Xq	IKBKG	NM_003639	153429255	153446455	10	inhibitor of kappa light polypeptide gene	Untereinheit der IKK-Signalsom Komplex-aktivierung	3
Xq	MTMR1	NM_003828	149612826	149684233	16	myotubularin-related protein 1	unbekannt	4
Xq	NSDHL	NM_015922	151750166	151788563	8	NAD(P) dependent steroid dehydrogenase-like	Enzym der Cholesterolsynthese	3
Xq	OPN1LW	NM_020061	153062938	153077701	6	opsin 1 (cone pigments), long-wave-sensitive	Visuelles Protein Opsin	3
Xq	OPN1MW	NM_000513	153101342	153114834	6	opsin 1 (cone pigments), medium-wave-sensitive	Visuelles Protein Opsin	3
Xq	OPN1MW	NM_000513	153138460	153151952	6	opsin 1 (cone pigments), medium-wave-sensitive	Visuelles Protein Opsin	3
Xq	PASD1	NM_173493	150482749	150595865	16	PAS domain containing 1	unbekannt	4
Xq	PDZD4	NM_032512	152720816	152749197	8	PDZ domain containing 4	unbekannt	4
Xq	PLXNA3	NM_017514	153339816	153355179	33	plexin A3	Protein der neuronalen und epithelialen Entwicklung	3
Xq	PLXNB3	NM_005393	152682904	152699134	36	plexin B3	Rezeptor für SEMA5A, Rolle in Axonleitung, invasivem Wachstum und Zellmigration	3
Xq	PNCK	NM_001039582	152588416	152592974	12	pregnancy upregulated non-ubiquitously expressed	Calcium/Caldmodulin-abhängige Proteinkinase in einer Calcium-ausgelösten Signalkaskade	3
Xq	PNMA3	NM_013364	151975450	151979478	3	paraneoplastic cancer-testis-brain antigen	Tumorentigen, Funktion unbekannt	4
Xq	PNMA5	NM_052926	151908026	151911414	2	paraneoplastic antigen like 5	unbekannt	4
Xq	PNMA6A	NM_032882	151991520	151994057	2	paraneoplastic antigen like 6A	unbekannt	4
Xq	PRRG3	NM_024082	150617434	150620669	4	proline rich Gla (G-carboxyglutamic acid) 3	unbekannt	4
Xq	RAB39B	NM_171998	154140719	154147046	2	RAB39B, member RAS oncogene family	Protein der RAB-Familie. Mögliche Rolle im vesiculären Trafficking	3
Xq	RENBP	NM_002910	152853916	152863426	11	renin binding protein	GlcNAc 2-Epimerase. Bindet und inhibiert Renin	3
Xq	RPL10	NM_006013	153279911	153283874	7	ribosomal protein L10	Ribosomales Protein der 60S Untereinheit	3
Xq	SLC10A3	NM_019848	153368841	153372189	2	solute carrier family 10, member 3	unbekannt	4
Xq	SLC6A8	NM_005629	152606585	152615234	13	solute carrier family 6	Aufnahme von Kreatin in Muskel und Gehirn	3
Xq	SLITRK2	NM_032539	144707041	144715052	5	SLIT and NTRK-like family, member 2	neuronaler Modulator	3
Xq	SLITRK4	NM_173078	142543607	142550685	2	slit and trk like 4 protein	neuronaler Modulator	3
Xq	SPANXA1	NM_013453	140499461	140500526	2	sperm protein associated with the nucleus, X	unbekannt	4
Xq	SPANXA2	NM_145662	140163261	140500526	2	sperm protein associated with the nucleus, X	unbekannt	4
Xq	SPANXE	NM_145665	140613233	140614321	2	sperm protein associated with the nucleus, X	unbekannt	4
Xq	SPANXN1	NM_001009614	144136039	144145324	2	SPANX-N1 protein	unbekannt	4
Xq	SPANXN2	NM_001009615	142622720	142632182	2	SPANX-N2 protein	unbekannt	4
Xq	SPANXN3	NM_001009609	142424229	142432973	2	SPANX-N3 protein	unbekannt	4
Xq	SPANXN4	NM_001009613	141941466	141949698	2	SPANX family, member N4	unbekannt	4
Xq	SPRY3	NM_005840	154650644	154665311	2	sprouty homolog 3	Antagonist des Fibroblasten-Wachstumsfaktor Signalwegs, negative Modulation der respiratorischen Organogenese	3
Xq	SRPK3	NM_014370	152699703	152704381	15	serine/threonine kinase 23	Proteinkinase	4
Xq	SSR4	NM_006280	152713287	152717148	6	signal sequence receptor, delta	Teil eines Komplexes, der in der Retention von ER-Proteinen wirkt	3
Xq	SYBL1	NM_005638	154764212	154826614	8	synaptobrevin-like 1	unbekannt	4
Xq	TAZ	NM_000116	153293070	153303257	11	tafazzin isoform 1	Cardiolipin-Metabolismus	3
Xq	TAZ	NM_181311	153293070	153303257	10	tafazzin isoform 2	Cardiolipin-Metabolismus	3
Xq	TAZ	NM_181312	153293070	153303257	10	tafazzin isoform 3	Cardiolipin-Metabolismus	3
Xq	TAZ	NM_181313	153293070	153303257	9	tafazzin isoform 4	Cardiolipin-Metabolismus	3
Xq	TAZ	NM_181314	153293070	153303257	10	tafazzin isoform 5	Cardiolipin-Metabolismus	3
Xq	TEX28	NM_001586	153152125	153176632	5	testis expressed 28	unbekannt	4
Xq	TKF1L	NM_012253	153177344	153211894	13	transketolase-like 1	Transketolase	4
Xq	TMEM185A	NM_032508	148486014	148521375	7	family with sequence similarity 11, member A	unbekannt	4
Xq	TMEM187	NM_003492	152891434	152901834	2	transmembrane protein 187	unbekannt	4



Tabelle 4: Gene in der Kandidatenregion 17q21.33-q25.1 (46654293-70734460)

Kandidaten region	GenSymbol	RefSeq- Nummer	Transkription Start	Transkription Ende	Exone	Name (englisch)	kurze Funktionsbeschreibung	Rang
17q	Sep 04	NM_004574	53952614	53961662	12	septin 4 isoform 1	Teil eines Komplexes mit Rolle in Zytokinese, pro-apoptotisches Protein	2
17q	Sep 04	NM_080415	53952614	53964410	10	septin 4 isoform 2	Teil eines Komplexes mit Rolle in Zytokinese, pro-apoptotisches Protein	2
17q	Sep 04	NM_080416	53952614	53964410	12	septin 4 isoform 3	Teil eines Komplexes mit Rolle in Zytokinese, pro-apoptotisches Protein	2
17q	ABCI	NM_022070	55475333	55511074	20	ATP binding cassette transporter 1	Anionentransporter, Einfluss auf Cholesterol-transport	3
17q	ABCA10	NM_080282	64655743	64732551	40	ATP-binding cassette, sub-family A member 10	Transporter mit mögl. Rolle in Lipidhomeostase der Makrophagen	3
17q	ABCA5	NM_018672	64754385	64823369	38	ATP-binding cassette, sub-family A member 5	mögl. Rolle im Autolysosomenumsatz	3
17q	ABCA5	NM_172232	64754385	64834885	39	ATP-binding cassette, sub-family A member 5	mögl. Rolle im Autolysosomenumsatz	3
17q	ABCA6	NM_080284	64586441	64649610	39	ATP-binding cassette, sub-family A member 6	Transporter mit mögl. Rolle in Lipidhomeostase der Makrophagen	3
17q	ABCA8	NM_007168	64375025	64463128	38	ATP-binding cassette, sub-family A member 8	lipophiler Drug-Transporter	3
17q	ABCA9	NM_080283	64482367	64568731	39	ATP-binding cassette, sub-family A member 9	mögl. Rolle in Differenzierung von Monozyten und Lipidhomeostase	3
17q	ACE	NM_000789	58908165	58928711	25	angiotensin I converting enzyme isoform 1	Umwandlung von Angiotensin I in Angiotensin II	3
17q	ACE	NM_152830	58915908	58928711	14	angiotensin I converting enzyme isoform 2	Umwandlung von Angiotensin I in Angiotensin II	3
17q	ACE	NM_152831	58915908	58952935	31	angiotensin I converting enzyme isoform 3	Umwandlung von Angiotensin I in Angiotensin II	3
17q	AKAP1	NM_003488	52517551	52553709	11	A-kinase anchor protein 1 precursor	Ankerprotein für Proteinkinase A	3
17q	AMZ2	NM_001033569	63755739	63764900	8	archaemetzincins-2 isoform 1	Zink-Metalloprotease	3
17q	AMZ2	NM_001033570	63755739	63764900	8	archaemetzincins-2 isoform 1	Zink-Metalloprotease	3
17q	AMZ2	NM_001033571	63755739	63764900	8	archaemetzincins-2 isoform 1	Zink-Metalloprotease	3
17q	AMZ2	NM_001033572	63755739	63764900	8	archaemetzincins-2 isoform 1	Zink-Metalloprotease	3
17q	AMZ2	NM_001033574	63755739	63764900	6	archaemetzincins-2 isoform 2	Zink-Metalloprotease	3
17q	AMZ2	NM_016627	63755739	63764900	7	archaemetzincins-2 isoform 1	Zink-Metalloprotease	3
17q	ANKFN1	NM_153228	51585834	51915006	17	ankyrin-repeat and fibronectin type III domain	unbekannt	4
17q	APOH	NM_000042	61638612	61655992	8	apolipoprotein H precursor	Bindung negativ geladener Substanzen	3
17q	APBP2	NM_006380	55875301	55988362	13	amyloid beta precursor protein-binding protein	Rolle im Transport und Processing des $\beta$ -Amyloid-Vorläuferproteins. Amplifiziert/Überexprimiert in Brustkrebs	2
17q	ARMC7	NM_024585	70617676	70637955	3	armadillo repeat containing 7	unbekannt	4
17q	ARSG	NM_014960	63814678	63928595	11	Arylsulfatase G	Sulfatase beteiligt an Hormonbiosynthese, Zellsignalmodulation und Degradation von Makromolekülen	3
17q	ATP5H	NM_001003785	70546549	70554669	5	ATP synthase, H+ transporting, mitochondrial F0	Untereinheit der F0-Komponente der mitochondrialen ATP-Synthase	3
17q	ATP5H	NM_006356	70546549	70554669	6	ATP synthase, H+ transporting, mitochondrial F0	Untereinheit der F0-Komponente der mitochondrialen ATP-Synthase	3
17q	AXIN2	NM_004655	60955144	60988227	11	axin 2	Regulation der Stabilität von $\beta$ -Catenin im Wnt-Signalweg. Mutationen in AXIN2 können zu einer onkogenen Aktivierung des Wnt-Signalweges führen.	2
17q	BCAS3	NM_017679	56110014	56824974	24	breast carcinoma amplified sequence 3	Überexprimiert in Brustkrebszelllinien, Funktion unbekannt	2
17q	BPTF	NM_004459	63252241	63410956	30	bromodomain PHD finger transcription factor	Histon-bindende Komponente des Nukleosomenbau-Faktors, der Transkription ermöglicht	3

17q	BPTF	NM_182641	63252241	63410956	28	bromodomain PHD finger transcription factor	Histon-bindende Komponente des Nukleosomenumbau-Faktors, der Transkription ermöglicht	3
17q	BRIP1	NM_032043	57114766	57295537	20	BRCA1 interacting protein C-terminal helcase 1	DNA-Helikase erforderlich für Erhalt der chromosomalen Stabilität, Rolle in Reparatur von DNA-Doppelstrangbrüchen durch homologe Rekombination, Assoziation mit BRCA1, bekanntes Brustkrebsuszeptibilitätsgen	1
17q	BZRAP1	NM_004758	53733592	53760477	32	peripheral benzodiazepine receptor-associated	unbekannt	4
17q	C17orf28	NM_030630	70458433	70480411	19	hypothetical protein LOC283987	mögl. Rolle in Krebsentstehung, Expressionsverlust in Brustkrebszelllinien	2
17q	C17orf47	NM_001038704	53973946	53976682	2	hypothetical protein LOC284083	unbekannt	4
17q	C17orf54	NM_182564	69257003	69336271	5	hypothetical protein LOC283982	unbekannt	4
17q	C17orf58	NM_181655	63420164	63420164	3	hypothetical protein LOC284018 isoform a	unbekannt	4
17q	C17orf58	NM_181656	63417678	63420164	3	hypothetical protein LOC284018 isoform b	unbekannt	4
17q	C17orf64	NM_181707	55854654	55863544	6	hypothetical protein LOC124773	unbekannt	4
17q	C17orf71	NM_018149	54642152	54647393	4	hypothetical protein LOC55181	unbekannt	4
17q	C17orf77	NM_152460	70092651	70101943	3	hypothetical protein LOC146723	unbekannt	4
17q	C17orf80	NM_017941	68740953	68755258	5	lung cancer-related protein 8	unbekannt	4
17q	C17orf82	NM_203425	56843893	56845423	1	hypothetical protein LOC388407	unbekannt	4
17q	CA10	NM_020178	47062682	47591086	9	carbonic anhydrase X	Rolle in ZNS-Entwicklung	3
17q	CA4	NM_000717	55582130	55591086	8	carbonic anhydrase IV precursor	Carboanhydrase, genaue Funktion nicht bekannt	4
17q	CACNG1	NM_000727	62483373	62483373	4	voltage-dependent calcium channel gamma-1	Untereinheit des L-Typ Calciumkanals	3
17q	CACNG4	NM_014405	62459980	62459980	4	voltage-dependent calcium channel gamma-4	Untereinheit Calciumkanal	3
17q	CACNG5	NM_014404	62303912	62311561	4	voltage-dependent calcium channel gamma-5	Untereinheit Calciumkanal	3
17q	CACNG5	NM_145811	62303912	62311819	5	voltage-dependent calcium channel gamma-5	Untereinheit Calciumkanal	3
17q	CCDC44	NM_016360	59031974	59039457	5	Coiled-coil domain-containing protein 44	unbekannt	4
17q	CCDC45	NM_138363	59933619	59964524	20	coiled-coil domain containing 45	unbekannt	4
17q	CCDC46	NM_001037325	61062119	61253063	7	coiled-coil domain containing 46 isoform b	unbekannt	4
17q	CCDC46	NM_145036	61062119	61618435	27	coiled-coil domain containing 46 isoform a	unbekannt	4
17q	CCDC47	NM_020198	59176341	59204668	13	coiled-coil domain containing 47	unbekannt	4
17q	CD300A	NM_007261	69974116	69992528	7	leukocyte membrane antigen	Rezeptor von Immunzellen	3
17q	CD300C	NM_006678	70048841	70053877	4	CD300C antigen	Immunzell-Antigen	3
17q	CD300E	NM_181449	70120306	70131396	4	immune receptor expressed on myeloid cells 2	Aktivierender Rezeptor in myeloidischen Zellen	3
17q	CD300LB	NM_174892	70028097	70039200	4	triggering receptor expressed on myeloid cells	unbekannt	4
17q	CD300LF	NM_139018	70202046	70220703	7	NK inhibitory receptor precursor	Rezeptor mit regulierender Funktion in myeloidischen Zellen	3
17q	CD79B	NM_000626	59359829	59363436	6	CD79B antigen isoform 1 precursor	Ig-β-Protein des B-Zell Antigenkomplexes	3
17q	CD79B	NM_001039933	59359829	59363436	6	CD79B antigen isoform 3 precursor	Ig-β-Protein des B-Zell Antigenkomplexes	3
17q	CD79B	NM_021602	59359829	59363436	5	CD79B antigen isoform 2 precursor	Ig-β-Protein des B-Zell Antigenkomplexes	3
17q	CDC42EP4	NM_012121	68791357	68819738	2	Cdc42 effector protein 4	Interaktion mit Proteinen der Rho-Familie GTPasen und Regulation der Organisation des Aktzytoskeletts	3
17q	C1DR2L	NM_014603	70495321	70513481	5	paraoneoplastic antigen	unbekannt	4
17q	CLTC	NM_004859	55051831	55129099	32	clathrin heavy chain 1	Clathrin-Untereinheit	3
17q	COG1	NM_018714	68700767	68716240	14	component of oligomeric golgi complex 1	Protein des Golgi-lokalisierten Komplexes, Rolle in Golgmorphologie und -funktion	3
17q	COIL	NM_004645	52370559	52393410	7	coilin	Komponente des Cajalkörpers	3
17q	COX11	NM_004375	50384263	50401053	5	COX11 homolog	unbekannt	4
17q	CSH1	NM_001317	59303103	59327719	5	chorionic sonatomamotropin hormone 1 isoform 1	gleich Wachstumshormon	3
17q	CSH1	NM_022640	59326004	59327719	4	chorionic sonatomamotropin hormone 1 isoform 2	gleich Wachstumshormon	3
17q	CSH1	NM_022641	59326004	59327719	3	chorionic sonatomamotropin hormone 1 isoform 3	gleich Wachstumshormon	3

17q	CSH2	NM_020991	59303103	59304821	5	chorionic somatomammotropin isoform 1	hormone 2	gleich Wachstumshormon	3
17q	CSH2	NM_022644	59303103	59304821	4	chorionic somatomammotropin isoform 2	hormone 2	gleich Wachstumshormon	3
17q	CSH2	NM_022645	59303103	59304821	3	chorionic somatomammotropin isoform 3	hormone 2	gleich Wachstumshormon	3
17q	CSHL1	NM_001318	59340696	59342350	4	chorionic somatomammotropin hormone-like 1		gleich Wachstumshormon	3
17q	CSHL1	NM_022579	59340696	59342350	5	chorionic somatomammotropin hormone-like 1		gleich Wachstumshormon	3
17q	CSHL1	NM_022580	59340696	59342350	4	chorionic somatomammotropin hormone-like 1		gleich Wachstumshormon	3
17q	CSHL1	NM_022581	59340696	59342350	5	chorionic somatomammotropin hormone-like 1		gleich Wachstumshormon	3
17q	CUEDC1	NM_017949	53295335	5335749	11	CUE domain-containing 1		unbekannt	4
17q	CYB561	NM_001017916	58863396	58877277	6	cytochrome b-561		Elektronen Transportprotein	3
17q	CYB561	NM_001017917	58863396	58877277	6	cytochrome b-561		Elektronen Transportprotein	3
17q	CYB561	NM_001915	58863396	58877454	6	cytochrome b-561		Elektronen Transportprotein	3
17q	DDX42	NM_007372	59205298	59250409	19	DEAD box polypeptide 42 protein		ATP-abhängige RNA helicase, Bilden von RNA-Strukturen	3
17q	DDX42	NM_203499	59205298	59250409	18	DEAD box polypeptide 42 protein		ATP-abhängige RNA helicase, Bilden von RNA-Strukturen	3
17q	DDX5	NM_004396	59926199	59982869	13	DEAD (Asp-Glu-Ala-Asp) box polypeptide 5		RNA-abhängige ATPase, mögl. Rolle in pre-mRNA Spleißen	3
17q	DGKE	NM_003647	52266551	52295922	12	diacylglycerol kinase epsilon		mögl. Rolle in Regulierung DAG-Signale oder Phospholipidsynthese	4
17q	DHX40	NM_024612	54997667	55040484	18	DEAH (Asp-Glu-Ala-His) box polypeptide 40		unbekannt	4
17q	DNAI2	NM_023036	69782006	69822618	14	dynein, axonemal, intermediate polypeptide 2		Protein des Dyneinkomplexes der respiratorischen Zilien	3
17q	DYNLL2	NM_080677	53515797	53521810	4	dynein light chain 2		beteiligt am Dynein-vermittelten intrazellulären Transport und Motilität	3
17q	EFCAB3	NM_173503	57811645	57847569	10	EF-hand calcium binding domain 3		unbekannt	4
17q	EPX	NM_000502	53625087	53636794	12	eosinophil peroxidase		Peroxidase	3
17q	ERN1	NM_001433	59474491	59561208	22	endoplasmic reticulum to nucleus signalling 1		Veränderung der Genexpression in Antwort auf ER-Stress, unfolded-protein response, mit Auslösung von Wachstumsarrest und Apoptose	2
17q	ERN1	NM_152461	59559843	59561208	1	endoplasmic reticulum to nucleus signalling 1		Veränderung der Genexpression in Antwort auf ER-Stress, unfolded-protein response, mit Auslösung von Wachstumsarrest und Apoptose	2
17q	FADS6	NM_178128	70385067	70401300	6	fatty acid desaturase domain family member 6		unbekannt	4
17q	FAM104A	NM_032837	68715099	68740013	3	hypothetical protein LOC84923		unbekannt	4
17q	FAM20A	NM_017565	64044859	64108690	11	family with sequence similarity 20, member A		unbekannt	4
17q	FAM33A	NM_182620	54542095	54587307	4	hypothetical protein LOC348235		Erhalt der Metaphaseplatte, Spindel-Kontrollpunkt Silencing	3
17q	FDXR	NM_004110	70370213	70380751	12	ferredoxin reductase isoform 2 precursor		Elektronentransferprotein, induziert p53-abh. durch DNA-Schaden in Zellen	2
17q	FDXR	NM_024417	70370213	70380751	12	ferredoxin reductase isoform 1 precursor		Elektronentransferprotein, induziert p53-abh. durch DNA-Schaden in Zellen	2
17q	FLJ42842	NM_001004335	46769074	46774931	3	CDNA FLJ42842 fis, clone BRCC2007034		unbekannt	4
17q	FTSJ3	NM_017647	59250524	59258188	21	FtsJ homolog 3		unbekannt	4
17q	GDPD1	NM_182569	54652615	54706655	10	glycerophosphodiester phosphodiesterase domain		unbekannt	4
17q	GH1	NM_000515	59348294	59349930	5	growth hormone 1 isoform 1		Wachstumshormon	3
17q	GH1	NM_022559	59348294	59349930	5	growth hormone 1 isoform 2		Wachstumshormon	3
17q	GH1	NM_022560	59348294	59349930	4	growth hormone 1 isoform 3		Wachstumshormon	3
17q	GH1	NM_022561	59348294	59349930	3	growth hormone 1 isoform 4		Wachstumshormon	3
17q	GH1	NM_022562	59348294	59349930	2	growth hormone 1 isoform 5		Wachstumshormon	3
17q	GH2	NM_002059	59311305	59312955	5	growth hormone 2 isoform 1		Wachstumshormon	3
17q	GH2	NM_022556	59311305	59312955	5	growth hormone 2 isoform 4		Wachstumshormon	3

17q	GH2	NM_022557	59311303	59312955	4	growth hormone 2 isoform 2	Wachstumshormon	3
17q	GH2	NM_022558	59311305	59312955	5	growth hormone 2 isoform 3	Wachstumshormon	3
17q	GNA13	NM_006572	60437294	60483216	4	guanine nucleotide binding protein (G protein)	Modulator/Umwandler verschiedener trans-membranöser Signalsysteme	4
17q	GPR142	NM_181790	69875239	69880334	4	G protein-coupled receptor 142	GPCR	4
17q	GPRC5C	NM_018653	69941430	69955163	4	G protein-coupled receptor family C, group 5	Retinoidsäure-induzierbarer GPCR	4
17q	GPRC5C	NM_022036	69939261	69955163	4	G protein-coupled receptor family C, group 5	Retinoidsäure-induzierbarer GPCR	4
17q	GRIN2C	NM_000835	70349762	70367602	13	N-methyl-D-aspartate receptor subunit 2C	NMDA-Rezeptor	3
17q	HELZ	NM_014877	62504527	62671774	33	helicase with zinc finger domain	RNA-Helicase mit Rolle im RNA-Metabolismus, Expressionsverlust spielt möglicherweise Rolle in Karzogenese	2
17q	HLF	NM_002126	50697319	50757425	4	hepatic leukemia factor	Transkriptionale Aktivator	4
17q	HN1	NM_001002032	70642938	70662370	4	hematological and neurological expressed 1	unbekannt	4
17q	HN1	NM_001002033	70642938	70662370	6	hematological and neurological expressed 1	unbekannt	4
17q	HN1	NM_016185	70642938	70662370	5	hematological and neurological expressed 1	unbekannt	4
17q	HSF5	NM_001080439	53852527	53920758	6	heat shock transcription factor family member 5	unbekannt	4
17q	ICAM2	NM_000873	59433687	59451699	6	intercellular adhesion molecule 2 precursor	Vermittlung wichtiger Interaktionen der Immunantwort	3
17q	ICT1	NM_001545	70520374	70528951	6	immature colon carcinoma transcript 1	Kolonkarzinom-Marker	4
17q	INTS2	NM_020748	57297512	57360159	25	integrator complex subunit 2	Protein des Integrator-Komplexes, der an snRNA U1 und U2 Transkription beteiligt ist	3
17q	KCNH6	NM_030779	58954426	58977334	14	potassium voltage-gated channel, subfamily H	Unterheit spannungsabhängiger Kaliumkanal	3
17q	KCNH6	NM_173092	58954426	58980070	15	potassium voltage-gated channel, subfamily H	Unterheit spannungsabhängiger Kaliumkanal	3
17q	KCNJ16	NM_018658	65583020	65643341	5	potassium inwardly-rectifying channel J16	Einwärtsrichter Kaliumkanal	3
17q	KCNJ16	NM_170741	65583020	65643341	4	potassium inwardly-rectifying channel J16	Einwärtsrichter Kaliumkanal	3
17q	KCNJ16	NM_170742	65612589	65643341	4	potassium inwardly-rectifying channel J16	Einwärtsrichter Kaliumkanal	3
17q	KCNJ2	NM_000891	65677270	65687778	2	potassium inwardly-rectifying channel J2	Einwärtsrichter Kaliumkanal	3
17q	KCTD2	NM_015353	70564873	70573576	6	potassium channel tetramerisation domain	unbekannt	4
17q	KIF19	NM_153209	69833930	69859053	12	kinesin family member 19	unbekannt	4
17q	KIP2B	NM_032559	49255260	49257572	1	kinesin protein	unbekannt	4
17q	KPNB2	NM_002266	63462309	63473432	11	karyopherin alpha 2	Nukleärer Proteinimport	3
17q	LIMD2	NM_030576	59126993	59131217	5	LIM domain containing 2	unbekannt	4
17q	LOC388419	NM_001080466	69864150	69869553	3	galectin-3-binding protein-like	unbekannt	4
17q	LOC51136	NM_016125	55384504	55396899	9	P/D016 protein	unbekannt	4
17q	LPO	NM_006151	53670846	53700878	13	lactoperoxidase	Enzym diverser Drüsen	3
17q	LRRC37A3	NM_199340	60280949	60345365	14	leucine rich repeat containing 37, member A3	unbekannt	4
17q	LYK5	NM_001003786	59133926	59172947	11	protein kinase LYK5 isoform 2	Essentielle Rolle in STK11-vermitteltem G1 Zellzyklusarrest. Regulation der Tumorsuppressoraktivität von STK11	2
17q	LYK5	NM_001003787	59133926	59172947	13	protein kinase LYK5 isoform 1	Essentielle Rolle in STK11-vermitteltem G1 Zellzyklusarrest. Regulation der Tumorsuppressoraktivität von STK11	2
17q	LYK5	NM_001003788	59133926	59172947	12	protein kinase LYK5 isoform 4	Essentielle Rolle in STK11-vermitteltem G1 Zellzyklusarrest. Regulation der Tumorsuppressoraktivität von STK11	2
17q	LYK5	NM_153335	59133926	59172947	9	protein kinase LYK5 isoform 3	Essentielle Rolle in STK11-vermitteltem G1 Zellzyklusarrest. Regulation der Tumorsuppressoraktivität von STK11	2
17q	MAP2K6	NM_002758	64922432	65050055	12	mitogen-activated protein kinase kinase 6	Rolle in Signaltransduktion, Aktivierung von p38 MAP Kinase, involviert in viele zelluläre Prozesse	4
17q	MAP3K3	NM_002401	59053532	59127402	16	mitogen-activated protein kinase kinase kinase 3	Rolle in Signaltransduktion	4
17q	MAP3K3	NM_203351	59053506	59127402	18	mitogen-activated protein kinase kinase kinase 3	Rolle in Signaltransduktion	4

17q	MEITL2A	NM_001005372	57854997	57881186	8	methytransferase like 2A	unbekannt	4
17q	MEITL2A	NM_181725	57854997	57881186	9	methytransferase like 2A	unbekannt	4
17q	MKS1	NM_017777	53637796	53651665	18	Meckel syndrome, type 1	unbekannt	4
17q	MMD	NM_012329	50824972	50851340	7	monocyte to macrophage	Mikrogliale Aktivierung, Ionenkanal in reifen Makrophagen	3
17q	MPO	NM_000250	53702215	53713295	12	myeloperoxidase	Enzym des Abwehrsystems	3
17q	MRC2	NM_006039	58058493	58124684	30	mannose receptor, C type 2	verschiedene Rezeptorfunktionen im zellulären Umsatz	3
17q	MRPS23	NM_016070	53271840	53282398	5	mitochondrial ribosomal protein S23	Ribosomales Protein der 28S Untereinheit	3
17q	MS12	NM_138962	52688929	53112298	14	musashi 2 isoform a	Regulation der Expression von mRNAs auf translationalem Level	3
17q	MS12	NM_170721	52688932	53065543	10	musashi 2 isoform b	Regulation der Expression von mRNAs auf translationalem Level	3
17q	MTMR4	NM_004687	53921896	53950211	19	myotubularin related protein 4	unbekannt	4
17q	NACAL	NM_199290	57022571	57023345	1	nascent-polypeptide-associated complex alpha	Verhinderung der Interaktion nicht-sekretorischer Polypeptide mit dem ER	3
17q	NAT9	NM_015654	70278280	70284065	7	N-acetyltransferase 9	unbekannt	4
17q	NOG	NM_005450	52026273	52027544	1	noggin precursor	Knoorpel Morphogenese, Embryogenese	3
17q	NOL1	NM_015462	63144522	63170728	18	nucleolar protein 11	unbekannt	4
17q	NT5C	NM_014595	70637917	70639472	5	5',3'-nucleotidase, cytosolic	Nukleotidase	3
17q	NUP85	NM_024844	70713191	70743449	19	nucleoporin 85	Protein des nukleären Porenkomplexes	3
17q	ORD1	NM_012374	53587513	53588446	1	olfactory receptor, family 4, subfamily D	Olfaktorischer Rezeptor	3
17q	ORD2	NM_001004707	53602015	53602939	1	olfactory receptor, family 4, subfamily D	Olfaktorischer Rezeptor	3
17q	OTOP2	NM_178160	70431964	70441235	7	otopetrin 2	unbekannt	4
17q	OTOP3	NM_178233	70443491	70457106	7	otopetrin 3	unbekannt	4
17q	PCIP	NM_021213	51183448	51209733	6	phosphatidylcholine transfer protein	Katalyseung des Phosphatidylcholin-Transfers zwischen Membranen	3
17q	PCTP	NM_021213	51183638	51209733	6	phosphatidylcholine transfer protein	Katalyseung des Phosphatidylcholin-Transfers zwischen Membranen	3
17q	PEGAM1	NM_000442	59753594	59817743	15	platelet/endothelial cell adhesion molecule	Zelladhäsionsmolekül	3
17q	PITPNC1	NM_012417	62804385	63120109	9	phosphatidylinositol transfer protein	Phosphatidylinositol-Transfer zwischen Membran-komponenten	3
17q	PITPNC1	NM_181671	62804385	63120109	10	phosphatidylinositol transfer protein	Phosphatidylinositol-Transfer zwischen Membran-komponenten	3
17q	POLG2	NM_007215	59904363	59923631	8	polymerase (DNA directed), gamma 2, accessory	Untereinheit der mitochondrialen Polymerase-Prozessivität	3
17q	PPM1D	NM_003620	56032335	56096818	6	protein phosphatase 1D	Feedback-Regulation des p38-p53 Signals, Suppressor der p53-vermittelten Transkription und Apoptose, Region häufig amplifiziert in Brustkrebs	2
17q	PPM1E	NM_014906	54188230	54417316	7	protein phosphatase 1E	Negativer Regulator des zellulären Stressreaktions-wegs	3
17q	PRKAR1A	NM_002734	64020137	64040505	11	cAMP-dependent protein kinase, regulatory	AMPK, Phosphorylierung von Zielproteinen	3
17q	PRKAR1A	NM_212471	64019704	64040505	11	cAMP-dependent protein kinase, regulatory	AMPK, Phosphorylierung von Zielproteinen	3
17q	PRKAR1A	NM_212472	64020137	64040505	11	cAMP-dependent protein kinase, regulatory	AMPK, Phosphorylierung von Zielproteinen	3
17q	PRKCA	NM_002737	61729357	62237324	17	protein kinase C, alpha	PKC, Phosphorylierung von Zielproteinen	3
17q	PRRI1	NM_018304	54587874	54633929	10	proline rich 11	unbekannt	4
17q	PSMC5	NM_002805	59258541	59263119	12	proteasome 26S ATPase subunit 5	26S Proteasomuntereinheit, Degradation ubiquitinerter Proteine	3
17q	PSMD12	NM_002816	62767080	62793183	11	proteasome 26S non-ATPase subunit 12	Regulatorische Untereinheit des 26S Proteasoms	3
17q	PTRH2	NM_001015509	55129448	55139638	3	Bcl-2 inhibitor of transcription isoform a	Fördert Caspase-unabhängige Apoptose durch Regulierung der Funktion transkriptionaler Regulatoren	2
17q	PTRH2	NM_016077	55129448	55139638	2	Bcl-2 inhibitor of transcription isoform b	Fördert Caspase-unabhängige Apoptose durch Regulierung der Funktion transkriptionaler Regulatoren	2

17q	RAB37	NM_001006638	70244955	70254394	9	RAB37, member RAS oncogene family isoform 2	Signaltransduktionsprotein, Rolle in Polarisationsprozessen in Epithelzellen	3
17q	RAB37	NM_175738	70178864	70255069	10	RAB37, member RAS oncogene family isoform 3	Signaltransduktionsprotein, Rolle in Polarisationsprozessen in Epithelzellen	3
17q	RAD51C*	NM_002876	54124961	54127694	2	RAD51 homolog C isoform 2	Reparatur durch homologe Rekombination von DNA-Doppelstrangbrüchen, potenzieller Tumorsuppressor; Region häufig amplifiziert in Brustkrebs	1
17q	RAD51C*	NM_058216	54124961	54166691	9	RAD51 homolog C isoform 1	Reparatur durch homologe Rekombination von DNA-Doppelstrangbrüchen, potenzieller Tumorsuppressor; Region häufig amplifiziert in Brustkrebs	1
17q	RGS9	NM_001081955	60564010	60654283	19	regulator of G-protein signalling 9 isoform 2	Inhibition der Signaltransduktion, Rolle in Phototransduktion	3
17q	RGS9	NM_003835	60564010	60654283	19	regulator of G-protein signalling 9 isoform 1	Inhibition der Signaltransduktion, Rolle in Phototransduktion	3
17q	RNF190	NM_152598	58132407	58239427	11	ring finger protein 190	Protein der Ubiquitination	3
17q	RNF43	NM_000999	53786036	53849893	9	ring finger protein 43	unbekannt	4
17q	RPL38	NM_000999	69711389	69717614	5	ribosomal protein L38	Ribosomales Protein der 60S Untereinheit	3
17q	RPL38	NM_001035258	69711389	69717614	5	ribosomal protein L38	Ribosomales Protein der 60S Untereinheit	3
17q	RPS6KB1	NM_003161	55325224	55382568	15	ribosomal protein S6 kinase, 70kDa, polypeptide	Kinase des ribosomalen Proteins S6, erhöhte Proteinsynthese und Zellproliferation, Amplifikation/Überexpression häufig in Brustkrebs	2
17q	SCN4A	NM_000334	59389645	59404010	24	voltage-gated sodium channel type 4 alpha	Spannungsabhängiger Natriumkanal	3
17q	SCPEP1	NM_021626	52410487	52439123	13	serine carboxypeptidase 1 precursor protein	unbekannt	4
17q	SDK2	NM_019064	68845406	68945808	39	sidekick 2	Zelladhäsionsprotein der neuronalen Entwicklung	3
17q	SFRS1	NM_001078166	53433278	53439706	3	splicing factor, arginine/serine-rich 1 isoform	Aktivierung/Repression pre-mRNA Spleißen	3
17q	SFRS1	NM_006924	53433278	53439706	4	splicing factor, arginine/serine-rich 1 isoform	Aktivierung/Repression pre-mRNA Spleißen	3
17q	SFC16A5	NM_004695	70595649	70613843	7	solute carrier family 16, member 5	Monocarboxylat Transporter	3
17q	SFC16A6	NM_004694	63775932	63799000	6	solute carrier family 16, member 6	Monocarboxylat Transporter	3
17q	SLC39A11	NM_139177	68153682	68600427	10	solute carrier family 39 (metal ion)	Metallionen Transporter	3
17q	SLC9A3R1	NM_004262	70256378	70277089	6	solute carrier family 9 (sodium/hydrogen)	Verbindung Plasmamembranproteine mit Zytoskelett, Interaktion mit Tumorsuppressoren SYK und MERLIN, mögl. Tumorsuppressoren in Brustkrebs	1
17q	SMARCD2	NM_003077	59263175	59274083	13	SWI/SNF-related matrix-associated	Protein des ATP-abh. Chromatin Remodelingkomplexes, Regulation der Transkription	3
17q	SMURF2	NM_022739	59971196	60088848	19	SMAD specific E3 ubiquitin protein ligase 2	E3 Ubiquitin-Protein Ligase, Protein-Degradation	3
17q	SOX9	NM_000346	67628755	67634155	3	transcription factor SOX9	Transkriptionsfaktor; Rolle in Skelettenwicklung, Chondrogenese	3
17q	SSTR2	NM_001050	68672754	68679657	2	somatostatin receptor 2	Somatostatin Rezeptor	3
17q	STXBP4	NM_178509	50401124	50596448	18	syntaxin binding protein 4	Translokation von Transportvesikeln, mögl. Rolle in Regulation der Insulinausschüttung	3
17q	SUMO2	NM_001005849	70675419	70690693	3	SM/T3 suppressor of mif two 3 homolog 2 isoform b	mögl. Funktion als Antagonist von Ubiquitin in Degradation, in Kerntransport, DNA-Replikation, Reparatur, Mitose und Signaltransduktion	2
17q	SUMO2	NM_006937	70675419	70690693	4	SM/T3 suppressor of mif two 3 homolog 2 isoform a	mögl. Funktion als Antagonist von Ubiquitin in Degradation, in Kerntransport, DNA-Replikation, Reparatur, Mitose und Signaltransduktion	2
17q	SUPT4H1	NM_003168	53777537	53784562	5	suppressor of Typ 4 homolog 1	Protein der Regulation des mRNA-Processing und transkriptionalen Elongation der RNA-Polymerase II	3
17q	TBX2	NM_005994	56832038	56841609	7	T-box 2	Transkriptionale Regulation, Mesoderm-differenzierung, Region häufig Amplifiziert/Überexprimiert in Brustkrebs	2
17q	TBX4	NM_018488	56888588	56916446	8	T-box 4	Transkriptionale Regulation, Mesoderm-differenzierung	3



17q	TEX14	NM_031272	53989037	54124415	32	testis expressed sequence 14 isoform b	Protein für Spermatogenese und männliche Fertilität	3
17q	TEX14	NM_198393	53989037	54124415	33	testis expressed sequence 14 isoform a	Protein für Spermatogenese und männliche Fertilität	3
17q	TEX2	NM_018469	59578526	59694385	12	testis expressed sequence 2	unbekannt	4
17q	THRAP1	NM_005121	57374747	57497425	30	mediator of RNA polymerase II transcription	Protein eines transkriptionalen Aktivations-komplexes	3
17q	TLK2	NM_006852	57910135	58044759	22	tousled-like kinase 2	Regulation des Chromatin-Assemblys	3
17q	TMEM100	NM_018286	51151988	51155141	2	hypothetical protein LOC55273	unbekannt	4
17q	TMEM104	NM_017728	70284268	70346215	10	hypothetical protein LOC54868	unbekannt	4
17q	TMEM49	NM_030938	55139810	55273235	12	transmembrane protein 49	Fördert Stress-induzierte Vakuolisierung, Zelltod	3
17q	TOM1L1	NM_005486	50333202	50394309	16	target of myb1-like 1	Protein der Signaltransduktion	3
17q	TRIM25	NM_005082	52320268	52346408	9	tripartite motif-containing 25	mögl. Transkriptionsfaktor, Upregulation der Expression durch Estrogene, onkogene Rolle in Brustkrebsprogression	2
17q	TRIM37	NM_001005207	54414791	54539011	25	tripartite motif-containing 37 protein	unbekannt	4
17q	TRIM37	NM_015294	54430346	54539011	24	tripartite motif-containing 37 protein	unbekannt	4
17q	TTYH2	NM_032646	69721290	69769752	14	twenty 2 isoform 1	Chlorid Kanal	3
17q	TTYH2	NM_052869	69756099	69769752	8	twenty 2 isoform 2	Chlorid Kanal	3
17q	TUBD1	NM_016261	55291632	55325078	9	delta-tubulin	Assoziiert mit Spermatidmanschette	3
17q	USH1G	NM_173477	70423770	70430946	3	Usher syndrome 1G protein	Rolle im Hörvorgang	3
17q	USP32	NM_032582	55609472	55824368	34	ubiquitin specific protease 32	unbekannt	4
17q	UTP18	NM_016001	46692895	46730291	14	UTP18, small subunit processome component	nukleolares Processing der pre-18S rRNA	4
17q	VEZF1	NM_007146	53403908	53420614	6	zinc finger protein 161	Transkriptionsfaktor	4
17q	WDR68	NM_005828	58981553	59025373	8	WD-repeat protein	unbekannt	4
17q	WIP1	NM_017983	63929016	63965210	14	WD40 repeat protein Interacting with yippee-like 2	mögl. Rolle in Autophagie, Protein-Trafficking	4
17q	YPEL2	NM_001005404	54763834	54833877	5	yippee-like 2	unbekannt	4

Tabelle 5: Gene in der Kandidatenregion 1q23.1-q23.3 (156199127-160592447)

Kandidaten region	GenSymbol	RefSeq- Nummer	Transkription Start	Transkription Ende	Exone	Name (englisch)	kurze Funktionsbeschreibung	Rang
1q	ADAM/TSS4	NM_005099	159426161	159435468	9	ADAM metalloproteinase with thrombospondin type 1	Degradation von Aggrakan und Brevikan	3
1q	AIM2	NM_004833	157298898	157313271	6	absent in melanoma 2	mögl. Rolle in tumorigenischen Reversion und Kontrolle der Proliferation. Expression unterdrückt Proliferation und Tumorigenität in Brustkrebszelllinien	2
1q	APCS	NM_001639	157824239	157825285	2	serum amyloid P component precursor	Glykoprotein, mögl. Chaperon, Kontrolle der Chromatindegradation	3
1q	APOA2	NM_001643	159458706	159460042	4	apolipoprotein A-II preprotein	HDL-Struktur und Metabolismus	3
1q	ARHGAP30	NM_001025598	159283355	159306384	12	Rho GTPase activating protein 30 isoform 1	Aktivierung von Rho GTPasen	3
1q	ARHGAP30	NM_181720	159283355	159306384	13	Rho GTPase activating protein 30 isoform 2	Aktivierung von Rho GTPasen	3
1q	ATF6	NM_007348	160002707	160195476	16	activating transcription factor 6	Transkriptionsfaktor der ER-Stress Antwort	3
1q	ATP1A2	NM_000702	158352171	158379998	23	Na <sup>+</sup> /K <sup>+</sup> -ATPase alpha 2 subunit preprotein	Untereinheit der Na <sup>+</sup> /K <sup>+</sup> -ATPase	3
1q	ATP1A4	NM_001001734	158413824	158423391	5	Na <sup>+</sup> /K <sup>+</sup> -ATPase alpha 4 subunit isoform 1	Untereinheit der Na <sup>+</sup> /K <sup>+</sup> -ATPase	3
1q	ATP1A4	NM_144699	158387986	158423391	22	Na <sup>+</sup> /K <sup>+</sup> -ATPase alpha 4 subunit isoform 2	Untereinheit der Na <sup>+</sup> /K <sup>+</sup> -ATPase	3
1q	B4GALT3	NM_003779	159407724	159413938	8	UDP-Gal:betaGlcNAc beta 1,4-	Enzym der Oligosaccharidsynthese	3
1q	C1orf192	NM_001013625	159601144	159604288	5	hypothetical protein LOC257177	unbekannt	4
1q	CADM3	NM_021189	157408043	157438426	10	immunoglobulin superfamily, member 4B	Rolle in Zell-Zelladhäsion	3
1q	CASQ1	NM_001231	158426988	158438300	11	skeletal muscle calsequestrin 1	Mitochondriales Protein, Calciumbindung und -speicherung	3
1q	CDC19	NM_012337	158108777	158136530	12	nasopharyngeal epithelium specific protein 1	unbekannt	4
1q	CD1A	NM_001763	156490550	156494882	6	CD1A antigen precursor	Presentation primärer Lipid- und Glykolipid-antigenen	3
1q	CD1B	NM_001764	156564363	156567945	6	CD1B antigen precursor	Presentation primärer Lipid- und Glykolipid-antigenen	3
1q	CD1C*	NM_001765	156526186	156531188	6	CD1C antigen precursor	Presentation primärer Lipid- und Glykolipid-antigenen	3
1q	CD1D	NM_001766	156416360	156422840	7	CD1D antigen precursor	Presentation primärer Lipid- und Glykolipid-antigenen	3
1q	CD1E	NM_001042583	156590163	156593967	6	CD1E antigen isoform b precursor	Presentation primärer Lipid- und Glykolipid-antigenen	3
1q	CD1E	NM_001042584	156590163	156593967	6	CD1E antigen isoform c precursor	Presentation primärer Lipid- und Glykolipid-antigenen	3
1q	CD1E	NM_001042585	156590163	156593967	6	CD1E antigen isoform d precursor	Presentation primärer Lipid- und Glykolipid-antigenen	3
1q	CD1E	NM_001042586	156590163	156593967	4	CD1E antigen isoform e precursor	Presentation primärer Lipid- und Glykolipid-antigenen	3
1q	CD1E	NM_001042587	156590163	156593967	4	CD1E antigen isoform f precursor	Presentation primärer Lipid- und Glykolipid-antigenen	3
1q	CD1E	NM_030893	156590163	156593967	6	CD1E antigen isoform a precursor	Presentation primärer Lipid- und Glykolipid-antigenen	3
1q	CD244	NM_016382	159066573	159099269	9	CD244 natural killer cell receptor 2B4	Aktivierung NK-Zell Rezeptor	3
1q	CD48	NM_001778	158915159	158948209	4	CD48 antigen (B-cell membrane protein)	Ligand für CD2, Interaktion zwischen Immunzellen	3
1q	CD84	NM_003874	158782418	158815387	7	CD84 antigen (leukocyte antigen)	Interaktion zwischen Immunzellen	3
1q	COPA	NM_004371	158525685	158580079	35	coatamer protein complex, subunit alpha	Protein des Coatamerkomplexes	3
1q	CRP	NM_000567	157948702	157951003	2	C-reactive protein, pentraxin-related	Funktionen in Immunabwehr	3
1q	DARC	NM_002036	157441133	157442914	2	Duffy blood group	nicht-spezifischer Rezeptor für Chemokine, verantwortlich für Duffy-Blutgruppensystem	3

1q	DEDD	NM_001039711	159357391	159369102	5	death effector domain-containing protein	Rolle in Caspasen-vermittelter Apoptose, Regulation der Degradation der Intermediär-filamente in Apoptose	2
1q	DEDD	NM_001039712	159357391	159368880	6	death effector domain-containing protein	Rolle in Caspasen-vermittelter Apoptose, Regulation der Degradation der Intermediär-filamente in Apoptose	2
1q	DEDD	NM_032998	159357391	159369102	6	death effector domain-containing protein	Rolle in Caspasen-vermittelter Apoptose, Regulation der Degradation der Intermediär-filamente in Apoptose	2
1q	DUSP12	NM_007240	159986204	159993576	6	dual specificity phosphatase 12	Negative Regulation von Proteinen der MAP-Familie	4
1q	DUSP23	NM_017823	158017382	158018957	3	dual specificity phosphatase 23	Phosphatase	4
1q	F11R	NM_144503	159232607	159275358	13	F11 receptor isoform a precursor	Regulierung von Tight Junctions	3
1q	F11R	NM_144504	159232607	159275358	12	F11 receptor isoform a precursor	Regulierung von Tight Junctions	3
1q	FCBR1A	NM_002001	157526129	157544638	6	Fc fragment of IgE, high affinity I, receptor	Untereinheit IgE-Rezeptor, Immunantwort	3
1q	FCBR1G	NM_004106	159451710	159455662	5	Fc fragment of IgE, high affinity I, receptor	Untereinheit IgE-Rezeptor, Immunantwort	3
1q	FCGR2A	NM_021642	159741843	159755594	7	Fc fragment of IgG, low affinity IIa, receptor	Untereinheit IgG-Rezeptor, Immunantwort	3
1q	FCGR2B	NM_001002273	159899563	159914575	7	Fc fragment of IgG, low affinity IIb, receptor	Untereinheit IgG-Rezeptor, Immunantwort	3
1q	FCGR2B	NM_001002274	159899563	159914575	7	Fc fragment of IgG, low affinity IIb, receptor	Untereinheit IgG-Rezeptor, Immunantwort	3
1q	FCGR2B	NM_001002275	159899563	159914575	8	Fc fragment of IgG, low affinity IIb, receptor	Untereinheit IgG-Rezeptor, Immunantwort	3
1q	FCGR2B	NM_001005412	159817761	159836656	8	Fc fragment of IgG, low affinity IIb, receptor	Untereinheit IgG-Rezeptor, Immunantwort	3
1q	FCGR2B	NM_004001	159899563	159914575	8	Fc fragment of IgG, low affinity IIb, receptor	Untereinheit IgG-Rezeptor, Immunantwort	3
1q	FCGR2C	NM_001005411	159817761	159836656	7	Fc fragment of IgG, low affinity IIc, receptor	Untereinheit IgG-Rezeptor, Immunantwort	3
1q	FCGR2C	NM_201563	159817761	159836656	8	Fc fragment of IgG, low affinity IIc, receptor	Untereinheit IgG-Rezeptor, Immunantwort	3
1q	FCGR3A	NM_000569	159778174	159786442	5	Fc fragment of IgG, low affinity IIIa, receptor	Untereinheit IgG-Rezeptor, Immunantwort	3
1q	FCGR3B	NM_000570	159859676	159867525	6	low affinity immunoglobulin gamma Fc region	Untereinheit IgG-Rezeptor, Immunantwort	3
1q	FCRL6	NM_001004310	158038833	158052075	10	Fc receptor-like 6	unbekannt	4
1q	FCRLA	NM_032738	159943385	159950766	5	Fc receptor-like and mucin-like 1	Untereinheit IgG-Rezeptor, Immunantwort	3
1q	FCRLB	NM_001002901	159959066	159964557	6	Fc receptor-like and mucin-like 2	Untereinheit IgG-Rezeptor, Immunantwort	3
1q	HSPA6	NM_002155	159760659	159763311	1	heat shock 70kDa protein 6 (HSP70B)	Chaperon	3
1q	IFI16	NM_005531	157246305	157291569	11	interferon, gamma-inducible protein 16	Kontrolle der Zellproliferation, Interaktion mit p53, Fb, BCR/Al. Reduzierte Expression assoziiert mit Brustkrebs	2
1q	IGSF8	NM_052868	158327753	158335032	7	immunoglobulin superfamily, member 8	Regulation der Proliferation, Differenzierung, Zellmotilität einzelner Zelltypen	3
1q	IGSF9	NM_020789	158163454	158182010	21	immunoglobulin superfamily, member 9	unbekannt	4
1q	ITLN1	NM_017625	159112953	159121584	8	intelectin	Rolle im Abwehrsystem gegen Mikroorganismen, Eisenmetabolismus	3
1q	ITLN2	NM_080878	159181439	159191213	8	intelectin 2	Rolle im Abwehrsystem	3
1q	KARCA1	NM_001007255	159334777	159336760	4	kelch/ankyrin repeat containing cyclin A1	Interaktion mit dem ZyklinA1-CDK2 Komplex	4
1q	KARCA1	NM_001007255	159334777	159336760	5	kelch/ankyrin repeat containing cyclin A1	Interaktion mit dem ZyklinA1-CDK2 Komplex	4
1q	KARCA1	NM_152366	159334777	159336760	4	kelch/ankyrin repeat containing cyclin A1	Interaktion mit dem ZyklinA1-CDK2 Komplex	4
1q	KCNJ10	NM_002241	158274656	158306585	2	potassium inwardly-rectifying channel, subfamily	Kaliumkanal	3
1q	KCNJ9	NM_004983	158317983	158325836	3	potassium inwardly-rectifying channel, subfamily	Einwärtsrichter Kaliumkanal	3
1q	KIRREL	NM_018240	156229686	156332470	13	kin of IRRE like	Rolle in normaler Entwicklung und Funktion der glomerulären Permeabilität	3
1q	LY9	NM_001033667	159032551	159039107	3	lymphocyte antigen 9 isoform b	Adhäsionsreaktion zwischen Immunzellen	3
1q	LY9	NM_002348	159032551	159064669	10	lymphocyte antigen 9 isoform a	Adhäsionsreaktion zwischen Immunzellen	3
1q	MNDA	NM_002432	157067791	157085894	7	myeloid cell nuclear differentiation antigen	Transkriptionsfaktor in myeloischer Zelllinie, Rolle in Antwort auf Interferon	3
1q	MPZ	NM_000530	159541150	159546377	6	myelin protein zero	Strukturprotein des peripheren Myelins	3
1q	NCSTN	NM_015331	158579086	158595366	17	nicastrin precursor	Kofaktor des gamma-Sekretase Komplexes	3

1q	NDUFS2	NM_004550	159435728	159450808	15	NADH dehydrogenase (ubiquinone) Fe-S protein 2	Komplex 1 der mitochondrialen Atmungskette	3
1q	NHLH1	NM_005598	158603497	158609226	2	nescient helix loop helix 1	Transkriptionsfaktor mit Kontrolle der Zelltyp Determination	4
1q	NTF1	NM_005600	159354514	159357481	7	nitrilase 1	Möglicher Tumorsuppressor	2
1q	NOS1AP	NM_014697	160306204	160604854	10	nitric oxide synthase 1 (neuronal) adaptor	Regulation in neuronaler NO-Synthese	3
1q	NRLI3	NM_001077469	159466079	159474624	9	constitutive androstane receptor isoform 6	Nukleärer Androstane-Rezeptor, Regulation der Transkription von Zielgenen	3
1q	NRLI3	NM_001077470	159466079	159474624	7	constitutive androstane receptor isoform 11	Nukleärer Androstane-Rezeptor, Regulation der Transkription von Zielgenen	3
1q	NRLI3	NM_001077471	159466079	159474624	8	constitutive androstane receptor isoform 5	Nukleärer Androstane-Rezeptor, Regulation der Transkription von Zielgenen	3
1q	NRLI3	NM_001077472	159466079	159474624	8	constitutive androstane receptor isoform 9	Nukleärer Androstane-Rezeptor, Regulation der Transkription von Zielgenen	3
1q	NRLI3	NM_001077473	159466079	159474624	8	constitutive androstane receptor isoform 12	Nukleärer Androstane-Rezeptor, Regulation der Transkription von Zielgenen	3
1q	NRLI3	NM_001077474	159466079	159474624	8	constitutive androstane receptor isoform 8	Nukleärer Androstane-Rezeptor, Regulation der Transkription von Zielgenen	3
1q	NRLI3	NM_001077475	159466079	159474624	7	constitutive androstane receptor isoform 15	Nukleärer Androstane-Rezeptor, Regulation der Transkription von Zielgenen	3
1q	NRLI3	NM_001077476	159466079	159474624	8	constitutive androstane receptor isoform 13	Nukleärer Androstane-Rezeptor, Regulation der Transkription von Zielgenen	3
1q	NRLI3	NM_001077477	159466079	159474624	8	constitutive androstane receptor isoform 14	Nukleärer Androstane-Rezeptor, Regulation der Transkription von Zielgenen	3
1q	NRLI3	NM_001077478	159466079	159474624	9	constitutive androstane receptor isoform 7	Nukleärer Androstane-Rezeptor, Regulation der Transkription von Zielgenen	3
1q	NRLI3	NM_001077479	159466079	159474624	8	constitutive androstane receptor isoform 10	Nukleärer Androstane-Rezeptor, Regulation der Transkription von Zielgenen	3
1q	NRLI3	NM_001077480	159466079	159474624	9	constitutive androstane receptor isoform 2	Nukleärer Androstane-Rezeptor, Regulation der Transkription von Zielgenen	3
1q	NRLI3	NM_001077481	159466079	159474624	8	constitutive androstane receptor isoform 4	Nukleärer Androstane-Rezeptor, Regulation der Transkription von Zielgenen	3
1q	NRLI3	NM_001077482	159466079	159474624	9	constitutive androstane receptor isoform 1	Nukleärer Androstane-Rezeptor, Regulation der Transkription von Zielgenen	3
1q	NRLI3	NM_005122	159466079	159474624	9	constitutive androstane receptor isoform 3	Nukleärer Androstane-Rezeptor, Regulation der Transkription von Zielgenen	3
1q	OLFML2B	NM_015441	160219605	160260268	8	olfactomedin-like 2B	unbekannt	4
1q	OR10J1	NM_012351	157676172	157677135	1	olfactory receptor, family 10, subfamily J	Olfaktorischer Rezeptor	3
1q	OR10J3	NM_001004467	157550083	157651073	1	olfactory receptor, family 10, subfamily J	Olfaktorischer Rezeptor	3
1q	OR10J5	NM_001004469	157771491	157772421	1	olfactory receptor, family 10, subfamily J	Olfaktorischer Rezeptor	3
1q	OR10K1	NM_001004473	156701975	156702917	1	olfactory receptor, family 10, subfamily K	Olfaktorischer Rezeptor	3
1q	OR10K2	NM_001004476	156656341	156657280	1	olfactory receptor, family 10, subfamily K	Olfaktorischer Rezeptor	3
1q	OR10R2	NM_001004472	156716291	156717299	1	olfactory receptor, family 10, subfamily R	Olfaktorischer Rezeptor	3
1q	OR10T2	NM_001004475	156634935	156635880	1	olfactory receptor, family 10, subfamily T	Olfaktorischer Rezeptor	3
1q	OR10X1	NM_001004478	156815332	156816313	1	olfactory receptor, family 10, subfamily X	Olfaktorischer Rezeptor	3
1q	OR10Z1	NM_001004478	156842852	156843794	1	olfactory receptor, family 10, subfamily Z	Olfaktorischer Rezeptor	3
1q	OR6K2	NM_001005279	156936091	156937066	1	olfactory receptor, family 6, subfamily K	Olfaktorischer Rezeptor	3
1q	OR6K3	NM_001005327	156953581	156954577	1	olfactory receptor, family 6, subfamily K	Olfaktorischer Rezeptor	3
1q	OR6K6	NM_001005184	156991229	156992261	1	olfactory receptor, family 6, subfamily K	Olfaktorischer Rezeptor	3
1q	OR6N1	NM_001005185	p157002157	157003096	1	olfactory receptor, family 6, subfamily N	Olfaktorischer Rezeptor	3
1q	OR6N2	NM_001005278	157013095	157014049	1	olfactory receptor, family 6, subfamily N	Olfaktorischer Rezeptor	3
1q	OR6Y1	NM_001005189	156783541	156784519	1	olfactory receptor, family 6, subfamily Y	Olfaktorischer Rezeptor	3
1q	PEA15	NM_003768	158441750	158451786	4	phosphoprotein enriched in astrocytes 15	Anti-apoptotisches Protein, Regulation des Glucosetransports. Überexprimiert in einigen Brustkrebszelllinien	2

1q	PEX19	NM_002857	158513225	158521555	8	peroxisomal biogenesis factor 19	Faktor der frühen Peroxisomenbiogenese, entfernt CDKN2A (Tumorsuppressor) aus Nukleus und verhindert dessen Interaktion mit MDM2, was die Degradation von TP53 zur Folge hat.	2
1q	PFDN2	NM_012394	159336969	159354490	4	prefoldin subunit 2	Untereinheit eines Chaperonkomplexes	3
1q	PIGM	NM_145167	158264085	158268407	1	phosphatidylinositol glycan anchor biosynthesis	Enzym der GPI-Anker Synthese	3
1q	PPOX	NM_000309	159402817	159407634	13	protoporphyrinogen oxidase	Enzym der Häm biosynthese	3
1q	PVR14	NM_030916	159308119	159252966	9	poliovirus receptor-related 4	Zell-Zelladhäsion, Brustkrebsantigen	4
1q	PYHIN1	NM_152501	157167965	157213467	9	pyrin and HIN domain family, member 1 alpha 1	mögliche Tumorsuppressoraktivität in Brustkrebs	1
1q	PYHIN1	NM_198928	157167965	157213467	9	pyrin and HIN domain family, member 1 alpha 2	mögliche Tumorsuppressoraktivität in Brustkrebs	1
1q	PYHIN1	NM_198929	157167965	157213467	8	pyrin and HIN domain family, member 1 beta 1	mögliche Tumorsuppressoraktivität in Brustkrebs	1
1q	PYHIN1	NM_198930	157167965	157213467	8	pyrin and HIN domain family, member 1 beta 2	mögliche Tumorsuppressoraktivität in Brustkrebs	1
1q	SDHC	NM_001035511	159550789	159601159	5	succinate dehydrogenase complex, subunit C	Untereinheit des Komplexes 2 der mitochondrialen Atmungskette, mögliche Rolle als Tumorsuppressor	2
1q	SDHC	NM_001035512	159550789	159601159	5	succinate dehydrogenase complex, subunit C	Untereinheit des Komplexes 2 der mitochondrialen Atmungskette, mögliche Rolle als Tumorsuppressor	2
1q	SDHC	NM_001035513	159550789	159601159	4	succinate dehydrogenase complex, subunit C	Untereinheit des Komplexes 2 der mitochondrialen Atmungskette, mögliche Rolle als Tumorsuppressor	2
1q	SDHC	NM_003001	159550789	159601159	6	succinate dehydrogenase complex, subunit C	Untereinheit des Komplexes 2 der mitochondrialen Atmungskette, mögliche Rolle als Tumorsuppressor	2
1q	SLAMF1	NM_003037	158846512	158883493	7	signaling lymphocytic activation molecule family	Ligand der bidirektionalen T- zu B-Zellstimulation	3
1q	SLAMF6	NM_052931	158721443	158759666	8	activating NK receptor precursor	Rezeptor der Immunantwort, NK-Zell Aktivierung	3
1q	SLAMF7	NM_021181	158975700	158991225	7	SLAM family member 7	NK-Zell Aktivierung: Lymphozytenadhäsion	3
1q	SLAMF8	NM_020125	158063224	158073906	5	B lymphocyte activator macrophage expressed	Immunantwort, Lymphozytenaktivierung	3
1q	SLAMF9	NM_033438	158187907	158190634	4	SLAM family member 9	Immunantwort, Lymphozytenaktivierung	3
1q	SPTA1	NM_003126	156847119	156923130	52	spectrin, alpha, erythrocytic 1	Verlinkung der Plasmamembran mit dem Aktinzytoskelett	3
1q	SUMO1	NM_003352	158553678	158554884	1	SMT3 suppressor of mif two 3 homolog 1 isoform a	Mögliche Rolle als Antagonist von Ubiquitin in Degradationsprozessen, Kerntransport, DNA-Replikation und Reparatur, Mitose und Signaltransduktion, Transkriptionsmodulation durch Unterdrückung BRCA1-vermittelter Aktivierung der Transkription	2
1q	TAGLN2	NM_003564	158154526	158161908	5	transgelin 2	unbekannt	4
1q	TOMM40L	NM_032174	159462456	159467031	10	translocase of outer mitochondrial membrane 40	unbekannt	4
1q	UPC1	NM_016406	159390172	159395270	6	Ufm1-conjugating enzyme 1	unbekannt	4
1q	USP1	NM_007122	159275664	159282381	11	upstream stimulatory factor 1 isoform 1	Transkriptionsfaktor	4
1q	USP1	NM_207005	159275664	159282381	11	upstream stimulatory factor 1 isoform 2	Transkriptionsfaktor	4
1q	USP21	NM_001014443	159395877	159402140	14	ubiquitin-specific protease 21	Abspaltung von Ubiquitin	3
1q	USP21	NM_012475	159395877	159402140	13	ubiquitin-specific protease 21	Abspaltung von Ubiquitin	3
1q	VANGL2	NM_020335	158636990	158665088	8	vang-like 2 (van gogh, Drosophila)	Rolle in Embryonalentwicklung	3
1q	VSI68	NM_001013661	158090729	158099071	7	V-set and immunoglobulin domain containing 8	unbekannt	4
1q	WDR42A	NM_015726	158452128	158498603	14	H326	unbekannt	4

Tabelle 6: Gene in der Kandidatenregion 20q12-q13.2 (40749905-49190813)

Kandidaten region	GenSymbol	RefSeq- Nummer	Transkription Start	Transkription Ende	Exone	Name (englisch)	kurze Funktionsbeschreibung	Rang
20q	ACOT8	NM_005469	43903767	43919442	6	peroxisomal acyl-CoA thioesterase 1 isoform a	Enzym in Peroxisomen, Fettsäurestoffwechsel	3
20q	ACOT8	NM_183386	43903767	43919442	5	peroxisomal acyl-CoA thioesterase 1 isoform c	Enzym in Peroxisomen, Fettsäurestoffwechsel	3
20q	ADA	NM_000022	42681576	42713790	12	adenosine deaminase	Enzym des Adenosinabbaus	3
20q	ADNP	NM_015339	48940289	48980934	5	activity-dependent neuroprotector	Vermittlung von VIP-assoziierten Effekten in normalem Wachstum und Tumorproliferation, Rolle in Embryogenese, Expressionssteigerung in Brustkrebsgewebe	2
20q	ADNP	NM_181442	48940289	48980934	4	activity-dependent neuroprotector	Vermittlung von VIP-assoziierten Effekten in normalem Wachstum und Tumorproliferation, Rolle in Embryogenese, Expressionssteigerung in Brustkrebsgewebe	2
20q	ARFGEF2	NM_006420	46971833	47083645	39	ADP-ribosylation factor guanine	Aktivierung von ADP-Ribosylierungs Faktoren durch beschleunigten Austausch von GDP mit GTP, beteiligt am Golgitransport	3
20q	B4GALT5	NM_004776	47682889	47763745	9	UDP-Gal:betaGlcNAc beta 1,4-	Enzym der Oligosaccharidsynthese	3
20q	BCAS4	NM_001010974	48844873	48927121	4	breast carcinoma amplified sequence 4 isoform c	Amplifiziert/Überexprimiert in vielen Brustkrebszelllinien, Funktion unbekannt	2
20q	BCAS4	NM_017843	48844873	48927121	6	breast carcinoma amplified sequence 4 isoform a	Amplifiziert/Überexprimiert in vielen Brustkrebszelllinien, Funktion unbekannt	2
20q	BCAS4	NM_198799	48844873	48927121	5	breast carcinoma amplified sequence 4 isoform b	Amplifiziert/Überexprimiert in vielen Brustkrebszelllinien, Funktion unbekannt	2
20q	C20orf10	NM_014477	43435933	43440371	5	TP53-target gene 5 protein	mögliche Rolle in TP53-vermittelten Signalwegen, negativer Regulator des Zellwachstums	2
20q	C20orf111	NM_016470	42258549	42272845	4	oxidative stress responsive 1	unbekannt	4
20q	C20orf121	NM_01039199	42537960	42556658	5	hypothetical protein LOC79183	unbekannt	4
20q	C20orf121	NM_024331	42537960	42556658	6	hypothetical protein LOC79183	unbekannt	4
20q	C20orf142	NM_001080472	42368610	42373303	2	hypothetical protein LOC128486	unbekannt	4
20q	C20orf165	NM_080608	43948536	43949645	2	chromosome 20 open reading frame 165	unbekannt	4
20q	C20orf175	NM_080829	48636032	48686833	22	hypothetical protein LOC140876	unbekannt	4
20q	C20orf67	NM_022104	43996723	44010069	17	phosphorylated CTD interacting factor 1	mögl. Rolle in transkriptionaler Elongation oder pre-mRNA Processing	4
20q	CD40	NM_001250	44180312	44191791	9	CD40 antigen isoform 1 precursor	Rezeptor der Immun- und inflammatorischen Reaktion	3
20q	CD40	NM_152854	44180312	44191791	8	CD40 antigen isoform 2 precursor	Rezeptor der Immun- und inflammatorischen Reaktion	3
20q	CDH22	NM_021248	44235782	44313741	11	cadherin 22 precursor	Zelladhäsionsprotein	3
20q	CEBPB	NM_005194	48240782	48242606	2	CCAAT/enhancer binding protein beta	Transkriptionsfaktor, Regulation von Genen der Immun- und inflammatorischen Reaktion	3
20q	CSE1L	NM_001316	47096244	47146893	25	CSE1 chromosome segregation 1-like protein	Vermittlung des Export von Importin-alpha, mögl. Rolle in Apoptose und Zellproliferation	2
20q	CTSA	NM_000308	43962997	43960865	15	cathepsin A precursor	Schutz lysosomaler Proteine	3
20q	DBNDD2	NM_001048221	43468084	43472661	4	SCF apoptosis response protein 1 isoform a	unbekannt	4
20q	DBNDD2	NM_001048222	43468084	43472661	4	SCF apoptosis response protein 1 isoform b	unbekannt	4
20q	DBNDD2	NM_001048223	43468650	43472661	4	SCF apoptosis response protein 1 isoform a	unbekannt	4
20q	DBNDD2	NM_001048224	43468650	43472661	4	SCF apoptosis response protein 1 isoform b	unbekannt	4
20q	DBNDD2	NM_001048225	43472661	43472661	6	SCF apoptosis response protein 1 isoform a	unbekannt	4
20q	DBNDD2	NM_001048226	43470328	43472661	5	SCF apoptosis response protein 1 isoform b	unbekannt	4
20q	DDX27	NM_017895	47269290	47294021	21	DEAD (Asp-Glu-Ala-Asp) box polypeptide 27	Putative RNA-Helikase	4
20q	DNTTIP1	NM_052951	43853982	43873473	13	terminal deoxynucleotidyltransferase interacting	unbekannt	4

20q	DPM1	NM_003859	48984811	49008467	9	dolichyl-phosphate mannosyltransferase	Mannosyltransferase	3
20q	ELMO2	NM_133171	44428096	44468678	22	engulfment and cell motility 2	Rolle in Zytoskeletturnordnung	3
20q	ELMO2	NM_182764	44428096	44468678	21	engulfment and cell motility 2	Rolle in Zytoskeletturnordnung	3
20q	EYA2	NM_172110	44956915	45250899	16	eyes absent 2 isoform c	Regulation der Transkription	4
20q	EYA2	NM_172111	45041528	45250899	16	eyes absent 2 isoform a	Regulation der Transkription	4
20q	EYA2	NM_172112	45041737	45250899	16	eyes absent 2 isoform a	Regulation der Transkription	4
20q	EYA2	NM_172113	45040280	45250899	16	eyes absent 2 isoform b	Regulation der Transkription	4
20q	FAM112A	NM_001008901	41788214	41789056	2	hypothetical protein LOC149699 isoform 2	unbekannt	4
20q	FAM112A	NM_176791	41788214	41789056	1	hypothetical protein LOC149699 isoform 1	unbekannt	4
20q	FLJ40606	NM_001039897	43992595	43997101	2	CDNA FLJ40606 fs, clone THYMU2011939	unbekannt	4
20q	GDAP1L1	NM_024034	42309321	42342427	6	ganglioside-induced differentiation-associated	unbekannt	4
20q	HNF4A	NM_000457	42463337	42493444	10	hepatocyte nuclear factor 4 alpha isoform b	Transkriptionsfaktor	3
20q	HNF4A	NM_001030003	42417854	42493444	10	hepatocyte nuclear factor 4 alpha isoform e	Transkriptionsfaktor	3
20q	HNF4A	NM_001030004	42417854	42493444	10	hepatocyte nuclear factor 4 alpha isoform f	Transkriptionsfaktor	3
20q	HNF4A	NM_175914	42417854	42493444	10	hepatocyte nuclear factor 4 alpha isoform d	Transkriptionsfaktor	3
20q	HNF4A	NM_175849	42463337	42493444	10	hepatocyte nuclear factor 4 alpha isoform a	Transkriptionsfaktor	3
20q	HNF4A	NM_178850	42463337	42486690	8	hepatocyte nuclear factor 4 alpha isoform c	Transkriptionsfaktor	3
20q	IFT52	NM_016004	41652992	41709276	14	intraflagellar transport 52 homolog	unbekannt	4
20q	JPH2	NM_020433	42173750	42249632	6	junctional protein 2 isoform 1	Komponente des junctionalen Komplexes zwischen Plasmamembran und ER/SR	3
20q	JPH2	NM_175913	42238870	42249632	2	junctional protein 2 isoform 2	Komponente des junctionalen Komplexes zwischen Plasmamembran und ER/SR	3
20q	KCNB1	NM_004975	47421911	47532590	2	potassium voltage-gated channel, Shab-related	Verzögerungsgleichrichter Kaliumkanal	3
20q	KCNB1	NM_002237	49053599	49060308	2	potassium voltage-gated channel, subfamily G	mögl. Untereinheit Kaliumkanal	3
20q	KCNB1	NM_172318	49058622	49073024	3	potassium voltage-gated channel, subfamily G	mögl. Untereinheit Kaliumkanal	3
20q	KCNK15	NM_022358	42807901	42813082	2	potassium family, subfamily K, member 15	mögl. Untereinheit Kaliumkanal	3
20q	KCNK15	NM_002251	43154363	43163167	5	potassium voltage-gated channel	Untereinheit Kaliumkanal	3
20q	Kua	NM_199129	48173714	48203667	6	ubiquitin-conjugating enzyme E2 variant 1	unbekannt	4
20q	Kua-UEV	NM_199203	48131067	48203667	8	ubiquitin-conjugating enzyme E2 Kua-UEV isoform	unbekannt	4
20q	L3MBTTL	NM_015478	41576466	41603949	19	l(3)mbl-like isoform I	Transkriptionaler Repressor	4
20q	L3MBTTL	NM_032107	41576466	41603949	18	l(3)mbl-like isoform II	Transkriptionaler Repressor	4
20q	LOC284751	NM_001025463	48342673	48361958	3	hypothetical protein LOC284751	unbekannt	4
20q	MATN4	NM_003833	43355500	43367670	9	matrilin 4 isoform 1 precursor	mögl. Rolle in Formation faserartiger Netzwerke in EZ-Matrix	3
20q	MATN4	NM_030590	43355500	43367670	8	matrilin 4 isoform 2 precursor	mögl. Rolle in Formation faserartiger Netzwerke in EZ-Matrix	3
20q	MATN4	NM_030592	43355500	43367670	7	matrilin 4 isoform 3 precursor	mögl. Rolle in Formation faserartiger Netzwerke in EZ-Matrix	3
20q	MMP9	NM_004994	44070953	44078607	13	matrix metalloproteinase 9 preproprotein	Rolle in lokaler Proteolyse der EZ-Matrix und Leukozytenmigration	3
20q	MOC53	NM_014484	49008769	49011227	1	molybdenum cofactor synthesis 3	Kofaktor der Molybdoenzyme	3
20q	MYBL2*	NM_002466	41729122	41778536	14	MYB-related protein B	Transkriptionsfaktor mit Rolle in Regulation von Zellüberleben, Proliferation und Differenzierung	4
20q	NCOA3	NM_006534	45564063	45719021	23	nuclear receptor coactivator 3 isoform b	Transkriptionsfaktor mit Rolle in Regulation von Zellüberleben, Proliferation und Differenzierung	2
20q	NCOA3	NM_181659	45564063	45719021	23	nuclear receptor coactivator 3 isoform a	Transkriptionsfaktor mit Rolle in Regulation von Zellüberleben, Proliferation und Differenzierung	2
20q	NCOA5	NM_020967	44123032	44151987	8	nuclear receptor coactivator 5	Koaktivator nukleärer Rezeptoren, veränderte Expression kann zur Entstehung Steroid-vermittelter MammaCA führen	4
20q	NEURL2	NM_080749	43950673	43953308	2	neutralized-like protein 2	Koaktivator nukleärer Rezeptoren, veränderte Expression kann zur Entstehung Steroid-vermittelter MammaCA führen	3
20q	PAR6B	NM_032521	48781487	48803685	3	PAR-6 beta	Rolle in asymmetrischer Zellteilung und Zellpolarisierungsprozessen	3

20q	PI3	NM_002638	43236911	43238599	3	elafin preproprotein	Elastaseinhibitor, Schutz vor Elastase-vermittelter Gewebelyse	3
20q	PIGT	NM_015937	43478137	43488298	12	phosphatidylinositol glycan anchor biosynthesis	Protein der GPI-Anker Biosynthese	3
20q	PKIG	NM_007066	42593849	42681092	4	cAMP-dependent protein kinase inhibitor gamma	Inhibitor der PKA	3
20q	PKIG	NM_181804	42593849	42681092	5	cAMP-dependent protein kinase inhibitor gamma	Inhibitor der PKA	3
20q	PKIG	NM_181805	42593849	42681092	6	cAMP-dependent protein kinase inhibitor gamma	Inhibitor der PKA	3
20q	PLTP	NM_006227	43960803	43974193	16	phospholipid transfer protein isoform a	Phospholipidtransferprotein, Rolle in Cholesterol-metabolismus	3
20q	PLTP	NM_182676	43960803	43974193	15	phospholipid transfer protein isoform b	Phospholipidtransferprotein, Rolle in Cholesterol-metabolismus	3
20q	PREX1	NM_020820	46674199	46877827	40	PREX1 protein	Guanin-Nukleotid Austauschfaktor	3
20q	PTGIS	NM_000961	47553817	47618114	10	prostaglandin I2 (prostaglandin) synthase	Enzym der Prostaglandinsynthese	3
20q	PTPN1	NM_002827	48560297	48634493	10	protein tyrosine phosphatase, non-receptor type	Regulation von Signaltransduktionskaskaden	4
20q	PTPRT	NM_007050	40134805	41251971	31	protein tyrosine phosphatase, receptor type, T	Regulation von Signaltransduktionskaskaden	4
20q	PTPRT	NM_133170	40134805	41251971	32	protein tyrosine phosphatase, receptor type, T	Regulation von Signaltransduktionskaskaden	4
20q	R3HDM1L	NM_178491	42399039	42413289	5	R3H domain (binds single-stranded nucleic acids)	mögl. Serinprotease-Inhibitor	4
20q	RBPJL	NM_014276	43368904	43379878	12	recombining binding protein L	mögl. Transkriptionsfaktor	4
20q	RIMS4	NM_182970	42813862	42872326	6	regulating synaptic membrane exocytosis 4	Regulation der Exozytose	3
20q	SDC4	NM_002999	43387342	43410478	5	syndecan 4 precursor	Proteoglykan mit Funktion als Rezeptor	4
20q	SEMG1	NM_003007	43269087	43271823	3	senenogelin I isoform a, preproprotein	Spermaprotein	3
20q	SEMG1	NM_198139	43269087	43271823	4	senenogelin I isoform b, preproprotein	Spermaprotein	3
20q	SEMG2	NM_003008	43283423	43286513	3	senenogelin II precursor	Spermaprotein	3
20q	SEPRNC3	NM_006811	42561314	42584140	10	tumor differentially expressed protein 1	unbekannt	4
20q	SEPRNC3	NM_198941	42568277	42584140	11	tumor differentially expressed protein 1	unbekannt	4
20q	SPRS6	NM_006275	41519931	41525655	7	arginine/serine-rich splicing factor 6	mRNA Spleißen	3
20q	SGK2	NM_016276	41628150	41647687	12	serum/glucocorticoid regulated kinase 2 isoform	Aktivierung von Kaliumkanälen	3
20q	SGK2	NM_170693	41621099	41647687	13	serum/glucocorticoid regulated kinase 2 isoform	Aktivierung von Kaliumkanälen	3
20q	SLC12A5	NM_020708	44091244	44122196	26	solute carrier family 12 member 5	K+/Cl- Kotransporter	3
20q	SLC13A3	NM_001011554	44619869	44746514	14	solute carrier family 13 member 3 isoform b	Natrium-Dicarboxylat Kotransporter	3
20q	SLC13A3	NM_022829	44619869	44713505	13	solute carrier family 13 member 3 isoform a	Natrium-Dicarboxylat Kotransporter	3
20q	SLC2A10	NM_030777	44771685	44798392	5	solute carrier family 2 member 10	Glukosetransporter	3
20q	SLC35C2	NM_015945	44411583	44426471	10	solute carrier family 35, member C2 isoform a	unbekannt	4
20q	SLC35C2	NM_173073	44411583	44426471	10	solute carrier family 35, member C2 isoform b	unbekannt	4
20q	SLC35C2	NM_173179	44411583	44426471	10	solute carrier family 35, member C2 isoform a	unbekannt	4
20q	SLC9A8	NM_015266	47862656	47942179	16	Na+/H+ exchanger isoform 8	NA+/H+-Austauscher, pH-Regulation, Signaltransduktion	3
20q	SLP1	NM_003064	43314292	43316620	4	secretory leukocyte peptidase inhibitor	Serinprotease-Inhibitor, Schutz der Gewebe vor Serinprotease	3
20q	SNAI1	NM_005985	48032933	48038827	3	snail 1 homolog	Transkriptionsfaktor, Rolle in Mesoderm-formierung	3
20q	SNX21	NM_001042632	43895876	43905321	4	sorting nexin 21 isoform c	unbekannt	4
20q	SNX21	NM_001042633	43895876	43905321	5	sorting nexin 21 isoform d	unbekannt	4
20q	SNX21	NM_033421	43895876	43905321	4	sorting nexin 21 isoform a	unbekannt	4
20q	SNX21	NM_152897	43895876	43905321	5	sorting nexin 21 isoform b	unbekannt	4
20q	SPATA2	NM_006038	47963335	47965475	3	spermatogenesis associated 2	mögl. Rolle in Spermatogenese	4
20q	SPINLW1	NM_020398	43602678	43609442	4	serine peptidase inhibitor-like, with Kunitz	Serinprotease-Inhibitor	4
20q	SPINLW1	NM_181502	43602678	43608430	3	serine peptidase inhibitor-like, with Kunitz	Serinprotease-Inhibitor	4



20q	SPINT4	NM_178455	43784401	43787749	3	serine peptidase inhibitor, Kunitz type 4	Serinprotease-Inhibitor	4
20q	STAU1	NM_001037328	47163282	47238311	14	staufen isoform c	Rolle in Positionierung der mRNA	3
20q	STAU1	NM_004602	47163282	47238311	14	staufen isoform a	Rolle in Positionierung der mRNA	3
20q	STAU1	NM_017452	47163282	47238311	13	staufen isoform a	Rolle in Positionierung der mRNA	3
20q	STAU1	NM_017453	47163282	47238311	14	staufen isoform b	Rolle in Positionierung der mRNA	3
20q	STAU1	NM_017454	47163282	47238311	12	staufen isoform a	Rolle in Positionierung der mRNA	3
20q	STK4	NM_006282	43028533	43142007	11	serine/threonine kinase 4	mögl. Tumorsuppressorfunktion in Kolorektalem Karzinom, Stress-aktivierte pro-apoptische Kinase	2
20q	SULF2	NM_018837	45719558	45848215	21	sulfatase 2 isoform a precursor	Modulierung der Heparansulfatwirkung	3
20q	SULF2	NM_198596	45719558	45848215	20	sulfatase 2 isoform b precursor	Modulierung der Heparansulfatwirkung	3
20q	SVS1	NM_033542	43425114	43431263	4	SVS1 Golgi-localized integral membrane protein	Protein trafficking	3
20q	TNNC2	NM_003279	43885261	43889360	6	fast skeletal muscle troponin C	Regulatorisches Protein der Muskelkontraktion	3
20q	TOMM34	NM_006809	43004184	43022528	7	translocase of outer mitochondrial membrane 34	Rolle in mitochondrialen Proteinimport	3
20q	TOX2	NM_032883	41978220	42131668	10	Granulosa cell HMG box protein 1	mögl. Transkriptionsaktivator	4
20q	TP53RK	NM_033550	44746410	44751683	2	p53-related protein kinase	Phosphorylierung von TP53, mögl. Teilfunktion in Aktivierung	2
20q	UBE2C*	NM_007019	43874661	43879003	6	ubiquitin-conjugating enzyme E2C isoform 1	Ubiquitinierung von Proteinen, Rolle in Abbau mitotischer Zykline und Zellzyklusprogression	3
20q	UBE2C*	NM_181799	43874661	43879003	6	ubiquitin-conjugating enzyme E2C isoform 2	Ubiquitinierung von Proteinen, Rolle in Abbau mitotischer Zykline und Zellzyklusprogression	3
20q	UBE2C*	NM_181800	43874661	43879003	5	ubiquitin-conjugating enzyme E2C isoform 3	Ubiquitinierung von Proteinen, Rolle in Abbau mitotischer Zykline und Zellzyklusprogression	3
20q	UBE2C*	NM_181801	43875060	43879003	6	ubiquitin-conjugating enzyme E2C isoform 4	Ubiquitinierung von Proteinen, Rolle in Abbau mitotischer Zykline und Zellzyklusprogression	3
20q	UBE2C*	NM_181802	43874661	43879003	5	ubiquitin-conjugating enzyme E2C isoform 4	Ubiquitinierung von Proteinen, Rolle in Abbau mitotischer Zykline und Zellzyklusprogression	3
20q	UBE2C*	NM_181803	43874661	43879003	4	ubiquitin-conjugating enzyme E2C isoform 5	Ubiquitinierung von Proteinen, Rolle in Abbau mitotischer Zykline und Zellzyklusprogression	3
20q	UBE2V1	NM_001032288	48131067	48163117	4	ubiquitin-conjugating enzyme E2 variant 1	Aktivierung der Transkription von Zielproteinen, Rolle in Zellzyklusverlauf, Differenzierung, DNA-Reparatur	2
20q	UBE2V1	NM_021988	48131067	48165901	4	ubiquitin-conjugating enzyme E2 variant 1	Aktivierung der Transkription von Zielproteinen, Rolle in Zellzyklusverlauf, Differenzierung, DNA-Reparatur	2
20q	UBE2V1	NM_022442	48131067	48165898	4	ubiquitin-conjugating enzyme E2 variant 1	Aktivierung der Transkription von Zielproteinen, Rolle in Zellzyklusverlauf, Differenzierung, DNA-Reparatur	2
20q	UBE2V1	NM_199144	48131067	48165901	5	ubiquitin-conjugating enzyme E2 variant 1	Aktivierung der Transkription von Zielproteinen, Rolle in Zellzyklusverlauf, Differenzierung, DNA-Reparatur	2
20q	WFDC10A	NM_080753	43691798	43693245	2	WAP four-disulfide core domain 10A precursor	mögl. Protease-Inhibitor	4
20q	WFDC10B	NM_172006	43746703	43767072	4	WAP four-disulfide core domain 10B isoform a	mögl. Protease-Inhibitor	4
20q	WFDC10B	NM_172131	43746703	43767072	3	WAP four-disulfide core domain 10B isoform b	mögl. Protease-Inhibitor	4
20q	WFDC11	NM_147197	43710615	43732292	5	WAP four-disulfide core domain 11 precursor	mögl. Protease-Inhibitor	4
20q	WFDC12	NM_080869	43185480	43186520	3	WAP four-disulfide core domain 12 precursor	mögl. Protease-Inhibitor	4
20q	WFDC13	NM_172005	43764068	43770870	4	WAP four-disulfide core domain 13 precursor	mögl. Protease-Inhibitor	4
20q	WFDC2	NM_006103	43531807	43543586	4	WAP four-disulfide core domain 2 isoform 1	mögl. Rolle in Spermiereifung	4
20q	WFDC2	NM_080734	43531807	43543586	3	WAP four-disulfide core domain 2 isoform 4	mögl. Rolle in Spermiereifung	4
20q	WFDC2	NM_080735	43541898	43543586	2	WAP four-disulfide core domain 2 isoform 5	mögl. Rolle in Spermiereifung	4
20q	WFDC3	NM_080736	43531807	43541802	3	WAP four-disulfide core domain 2 isoform 2	mögl. Rolle in Spermiereifung	4
20q	WFDC3	NM_080614	43836253	43853954	7	WAP four-disulfide core domain 3 precursor	mögl. Protease-Inhibitor	4
20q	WFDC3	NM_181522	43836253	43853954	5	WAP four-disulfide core domain 3 isoform 2	mögl. Protease-Inhibitor	4
20q	WFDC3	NM_181525	43836253	43853954	5	WAP four-disulfide core domain 3 isoform 2	mögl. Protease-Inhibitor	4

20q	WFDC3	NM_181530	43836253		3	WAP four-disulfide core domain 3 isoform 2	mögl. Protease-Inhibitor	4
20q	WFDC5	NM_145652	43171506		4	WAP four-disulfide core domain 5 precursor	mögl. Protease-Inhibitor	4
20q	WFDC6	NM_080827	43596249		3	WAP four-disulfide core domain 6 precursor	mögl. Protease-Inhibitor	4
20q	WFDC8	NM_130896	43613814		6	WAP four-disulfide core domain 8 precursor	mögl. Protease-Inhibitor	4
20q	WFDC8	NM_181510	43613204		7	WAP four-disulfide core domain 8 precursor	mögl. Protease-Inhibitor	4
20q	WFDC9	NM_147198	43669991		5	protease inhibitor WAP9	mögl. Protease-Inhibitor	4
20q	WISP2	NM_003881	42777298		4	WNT1 inducible signaling pathway protein 2	nachfolgendes Gen im WNT1 Signalweg,	2
20q	YWHAB	NM_003404	42947757		7	tyrosine 3-monooxygenase/tryptophan	Expression assoziiert mit Brustkrebs	4
20q	YWHAB	NM_139323	42947757		6	tyrosine 3-monooxygenase/tryptophan	Regulatorische Funktion in verschiedenen Signalwegen	4
20q	ZMYND8	NM_012408	45271787		22	zinc finger, MYND-type containing 8 isoform b	Regulatorische Funktion in verschiedenen Signalwegen	4
20q	ZMYND8	NM_183047	45271787		23	zinc finger, MYND-type containing 8 isoform a	unbekannt	4
20q	ZMYND8	NM_183048	45271787		22	zinc finger, MYND-type containing 8 isoform c	unbekannt	4
20q	ZNF313	NM_018683	47986320		6	zinc finger protein 313	mögl. Rolle in Spermatogenese	4
20q	ZNF334	NM_018102	44563113		5	zinc finger protein 334 isoform a	mögl. Rolle in Transkriptionsregulation	4
20q	ZNF334	NM_199441	44563113		6	zinc finger protein 334 isoform b	mögl. Rolle in Transkriptionsregulation	4
20q	ZNF335	NM_022095	44010698		28	zinc finger protein 335	Transkriptionsaktivierung	4
20q	ZNFX1	NM_021035	47295845		14	zinc finger, NFX1-type containing 1	unbekannt	4
20q	ZSWIM1	NM_080603	43943286		3	zinc finger, SWIM domain containing 1	unbekannt	4
20q	ZSWIM3	NM_080752	43919668		2	zinc finger, SWIM domain containing 3	unbekannt	4

Tabelle 7: Gene in der Kandidatenregion 7q36.2-q36.3 (153661756-156156741)

Kandidaten region	GenSymbol	RefSeq- Nummer	Transkription Start	Transkription Ende	Exone	Name (englisch)	kurze Funktionsbeschreibung	Rang
7q	DPP6	NM_001039350	153215351	154316928	26	dipeptidyl-peptidase 6 isoform 3	Regulierung spezifischer spannungsgesteuerter Kaliumkanäle	3
7q	DPP6	NM_001936	153633279	154316928	26	dipeptidyl-peptidase 6 isoform 2	Regulierung spezifischer spannungsgesteuerter Kaliumkanäle	3
7q	DPP6	NM_130797	153380709	154316928	26	dipeptidyl-peptidase 6 isoform 1	Regulierung spezifischer spannungsgesteuerter Kaliumkanäle	3
7q	EN2	NM_001427	154943584	154950287	2	engrailed homeobox 2	Rolle in Entwicklung der DNS, mögliches Onkogen in Brustkrebs	2
7q	HTR5A	NM_024012	154493478	154508392	2	5-hydroxytryptamine receptor 5A	Rezeptor für Serotonin	3
7q	INSIG1	NM_005542	154720475	154732879	6	insulin induced gene 1 isoform 1	Feedback-Kontrolle der Cholesterolsynthese	3
7q	INSIG1	NM_198336	154720475	154732879	7	insulin induced gene 1 isoform 2	Feedback-Kontrolle der Cholesterolsynthese	3
7q	INSIG1	NM_198337	154720475	154732879	4	insulin induced gene 1 isoform 3	Feedback-Kontrolle der Cholesterolsynthese	3
7q	PAXIP1*	NM_007349	154366332	154425615	21	PAX interacting protein 1	Rolle in Erhaltung der Genomstabilität, Chromatinkondensation und Mitoseprogression, Schlüsselkomponente der Reaktion auf DNA-Schaden	2
7q	PRR8	NM_053043	155158321	155231175	12	proline rich 8	unbekannt	4
7q	RBM33	NM_001008408	155232040	155266935	5	RNA binding motif protein 33	unbekannt	4
7q	RNF32	NM_030936	156126161	156162581	9	ring finger protein 32	mögl. Rolle in Spermatogenese	4
7q	SHH	NM_000193	155288318	155297728	3	sonic hedgehog preproprotein	Rolle in Embryonalentwicklung	3

Tabelle 8: Gene in der Kandidatenregion 8p23.3 (1-2117883)

Kandidaten region	GenSymbol	RefSeq- Nummer	Transkription Start	Transkription Ende	Exone	Name (englisch)	kurze Funktionsbeschreibung	Rang
8p	ARHGGEF10	NM_014629	1759555	1894214	29	Rho guanine nucleotide exchange factor 10	Bildung eines Komplex mit G-Proteinen und Stimulierung Rho-abh. Signale	3
8p	C8orf42	NM_175075	431645	485331	3	hypothetical protein LOC157695	unbekannt	4
8p	C8orf61	NM_001034061	1720280	1722142	1	hypothetical protein LOC619435	ER-Protein mit Rolle in Lipidmetabolismus	3
8p	CLN8	NM_018941	1699347	1719859	3	CLN8 protein	ER-Protein mit Rolle in Lipidmetabolismus	3
8p	DLGAP2	NM_004745	1436975	1644049	12	discs large-associated protein 2	Organisation der Synapsen, neuronales Zellsignal	3
8p	ERIC1	NM_207332	604199	671226	6	glutamate-rich 1	unbekannt	4
8p	FBXO25	NM_012173	346807	409875	9	F-box only protein 25 isoform 3	Komponente des E3 Ubiquitinligase Komplexes	3
8p	FBXO25	NM_183420	346807	409875	10	F-box only protein 25 isoform 2	Komponente des E3 Ubiquitinligase Komplexes	3
8p	FBXO25	NM_183421	346807	409875	11	F-box only protein 25 isoform 1	Komponente des E3 Ubiquitinligase Komplexes	3
8p	KBTD11	NM_014867	1909450	1942509	2	kelch repeat and BTB (POZ) domain containing 11	unbekannt	4
8p	LOC389607	NM_001013651	597191	601619	3	hypothetical protein LOC389607	unbekannt	4
8p	MYOM2	NM_003970	1980654	2080787	37	myomesin 2	Komponente der myofibrillären M-Bande	3
8p	OR4F21	NM_001005504	106085	107024	1	olfactory receptor, family 4, subfamily F	Olfaktorischer Rezeptor	3
8p	ZNF596	NM_001042415	172383	187339	6	zinc finger protein 596	mögl. Rolle in Regulation der Transkription	4
8p	ZNF596	NM_001042416	172383	187339	6	zinc finger protein 596	mögl. Rolle in Regulation der Transkription	4
8p	ZNF596	NM_173539	172199	187339	6	zinc finger protein 596	mögl. Rolle in Regulation der Transkription	4

Tabelle 9: Gene in der Kandidatenregion 13q32.1-q34 (96484285-114142980)

Kandidaten region	GenSymbol	RefSeq- Nummer	Transkription Start	Transkription Ende	Exone	Name (englisch)	kurze Funktionsbeschreibung	Rang
13q	ABHD13	NM_032859	107668763	107684604	2	abhydrolase domain containing 13	unbekannt	4
13q	ADPRHL1	NM_138430	113124260	113155540	7	ADP-ribosylhydrolase like 1 isoform 1	unbekannt	4
13q	ADPRHL1	NM_199162	113124260	113151458	7	ADP-ribosylhydrolase like 1 isoform 2	unbekannt	4
13q	ANKRD10	NM_017664	110328887	110865417	6	ankyrin repeat domain 10	unbekannt	4
13q	ARHGEF7	NM_003899	110604085	110756079	20	Rho guanine nucleotide exchange factor 7 isoform	Guannukleotid-Austauschfaktor für Rho-Proteine, Stimulierung Rho-abh. Signale	3
13q	ARHGEF7	NM_145735	110565782	110756079	19	Rho guanine nucleotide exchange factor 7 isoform	Guannukleotid-Austauschfaktor für Rho-Proteine, Stimulierung Rho-abh. Signale	3
13q	ARHGEF7	NM_145735	110565782	110756079	20	Rho guanine nucleotide exchange factor 7 isoform	Guannukleotid-Austauschfaktor für Rho-Proteine, Stimulierung Rho-abh. Signale	3
13q	ARHGEF7	NM_145735	110565782	110756079	18	Rho guanine nucleotide exchange factor 7 isoform	Guannukleotid-Austauschfaktor für Rho-Proteine, Stimulierung Rho-abh. Signale	3
13q	ATP11A	NM_015205	112392643	112589469	30	ATPase, Class VI, type 11A isoform a	Integrale Membran-ATPase, mögl. Rolle im Ionentransport	3
13q	ATP11A	NM_032189	112392643	112589469	29	ATPase, Class VI, type 11A isoform b	Integrale Membran-ATPase, mögl. Rolle im Ionentransport	3
13q	ATP4B	NM_000705	113351122	113360502	7	ATPase, H+/K+ exchanging, beta	Unterinheit Gastrische H+/K+-ATPase	3
13q	BIVM	NM_017693	102249399	102291882	11	basic, immunoglobulin-like variable motif	unbekannt	4
13q	C13orf16	NM_152324	110771015	110794595	6	hypothetical protein LOC121793	unbekannt	4
13q	C13orf27	NM_138779	102216466	102224143	6	hypothetical protein LOC93081	unbekannt	4
13q	C13orf28	NM_145248	112078669	112137002	7	hypothetical protein LOC122258	unbekannt	4
13q	C13orf8	NM_032436	114098116	114110899	3	hypothetical protein LOC283489	unbekannt	4
13q	CARS2	NM_024537	110091759	110156464	15	cysteine-tRNA synthetase 2	unbekannt	4
13q	CDC16	NM_001078645	114018463	114056252	18	anaphase-promoting complex subunit6	Komponente des Anaphase-Promoting Komplexes	3
13q	CDC16	NM_003903	114018463	114056252	19	anaphase-promoting complex subunit6	Komponente des Anaphase-Promoting Komplexes	3
13q	CLYBL	NM_138280	99056936	99359714	10	citrate lyase beta like	unbekannt	4
13q	CLYBL	NM_206808	99056936	99342825	9	citrate lyase beta like	unbekannt	4
13q	COL4A1	NM_001845	109599310	109757459	52	alpha 1 type IV collagen preproprotein	Strukturkomponente der Basalmembran	3
13q	COL4A2	NM_001846	109757631	109963374	48	alpha 2 type IV collagen preproprotein	Strukturkomponente der Basalmembran	3
13q	CUL4A	NM_001008895	112911931	112967393	21	cullin 4A isoform 1	Rolle in Ubiquitinierung und Degradation von Zielproteinen, wichtig für DNA-Reparatur, potenzielles Onkogen, Amplifikation/Überexpression in Brustkrebs	2
13q	CUL4A	NM_003589	112911086	112967393	21	cullin 4A isoform 2	Rolle in Ubiquitinierung und Degradation von Zielproteinen, wichtig für DNA-Reparatur, potenzielles Onkogen, Amplifikation/Überexpression in Brustkrebs	2
13q	DAOA	NM_172370	104916364	104941384	5	D-amino acid oxidase activator	Aktivierung von D-Aminosäure-Oxidase mit Wirkung auf NMDA-Aktivierung	3
13q	DCUN1D2	NM_001014283	113158134	113193024	7	DCN1, defective in cullin neddylation 1, domain containing 2	mögl. Rolle in Neddylation von Cullin-Komponenten des E3 Ubiquitinligase Komplexes	4
13q	DOCK9	NM_015296	98243744	98536661	56	dedicator of cytokinesis 9	Guannukleotid-Austauschfaktor, Aktivierung von CDC42	3
13q	DOCK9	NM_015296	98259606	98428339	48	dedicator of cytokinesis 9	Guannukleotid-Austauschfaktor, Aktivierung von CDC42	3
13q	EBI2	NM_004951	98744789	98757695	2	EBV-induced G protein-coupled receptor 2	unbekannt	4
13q	EFNB2	NM_004093	105940096	105985338	5	ephrin B2	mögl. Rolle in Entwicklung und Onkogenese verschiedener Gewebe	4

13q		ERCC5	NM_000123	102296174	102326346	15	XPG-complementing protein	DNA Endonuklease in DNA Exzisions-Reparatur	2
13q		F10	NM_000504	112825113	112851844	8	coagulation factor VII preproprotein	Faktor X der Blutgerinnung	3
13q		F7	NM_000131	112808105	112822996	9	coagulation factor VII isoform a precursor	Faktor VII der Blutgerinnung	3
13q		F7	NM_019616	112808105	112822996	8	coagulation factor VII isoform b precursor	Faktor VII der Blutgerinnung	3
13q		FAM70B	NM_182614	113599043	113651873	9	family with sequence similarity 70, member B	unbekannt	4
13q		FARP1	NM_001001715	97593434	97695559	3	FERM, RhoGEF and pleckstrin domain protein 1	mögl. Rho Guaninnukleotid-Austauschfaktor	4
13q		FARP1	NM_005766	97593434	97900024	27	FERM, RhoGEF and pleckstrin domain protein 1	mögl. Rho Guaninnukleotid-Austauschfaktor	4
13q		FGF14	NM_004115	101173035	101366996	5	fibroblast growth factor 14 isoform 1A	mögl. Funktion in ZNS-Entwicklung und Funktion	4
13q		FGF14	NM_175929	101173035	101852125	5	fibroblast growth factor 14 isoform 1B	mögl. Funktion in ZNS-Entwicklung und Funktion	4
13q		FLJ10154	NM_018011	105993675	106018483	3	hypothetical protein LOC55082	unbekannt	4
13q		FLJ10769	NM_018210	110066008	110900343	10	hypothetical protein LOC55739	unbekannt	4
13q		FLJ20443	NM_207440	112349358	112386812	3	hypothetical protein LOC400165	unbekannt	4
13q		GAS6	NM_000820	113546912	113590396	15	growth arrest-specific 6	Stimulation der Zellproliferation, Amplifikation/Überexpression in Brustkrebs	2
13q		GPRI8	NM_005292	98704971	98708841	3	G protein-coupled receptor 18	Rezeptor für N-Arachidonyl Glycin	3
13q		GRK1	NM_002929	113369597	113373973	3	rhodopsin kinase	Phosphorylierung von Rhodopsin und dessen Inaktivierung	3
13q		GRTP1	NM_024719	113026505	113066464	8	growth hormone regulated TBC protein 1	mögl. GTPase aktivierendes Protein	4
13q		ING1	NM_005537	110165359	110171422	2	inhibitor of growth family, member 1 isoform D	Tumorsuppressor	1
13q		ING1	NM_198217	110163083	110171422	2	inhibitor of growth family, member 1 isoform C	Tumorsuppressor	1
13q		ING1	NM_198218	110163610	110171422	2	inhibitor of growth family, member 1 isoform B	Tumorsuppressor	1
13q		ING1	NM_198219	110163625	110171422	2	inhibitor of growth family, member 1 isoform A	Tumorsuppressor	1
13q		IRS2	NM_003749	109204184	109236915	2	insulin receptor substrate 2	Signalmolekül der Vermittlung von Effekten des Insulins und anderer Zytokine	3
13q		ITGBL1	NM_004791	100902966	101166795	11	integrin, beta-like 1 (with EGF-like repeat)	unbekannt	4
13q		KDELC1	NM_024089	102234631	102249358	10	KDEL (Lys-Asp-Glu-Leu) containing 1	unbekannt	4
13q		LAMP1	NM_005561	112999469	113025742	9	lysosomal-associated membrane protein 1	Membranäres Glykoprotein, Präsentation von Kohlenhydratligenanden an Selektine, Amplifikation/Überexpression in Brustkrebs	2
13q		LIG4	NM_002312	107657792	107665131	2	DNA ligase IV	DNA-Ligase der Reparatur von DNA-Doppelstrangbrüchen durch nonhomologous end joining	2
13q		LIG4	NM_206937	107657792	107665883	3	DNA ligase IV	DNA-Ligase der Reparatur von DNA-Doppelstrangbrüchen durch nonhomologous end joining	2
13q		LIG4	NM_206937	107657792	107668717	2	DNA ligase IV	DNA-Ligase der Reparatur von DNA-Doppelstrangbrüchen durch nonhomologous end joining	2
13q		LOC196541	NM_001010977	102136097	102144870	4	hypothetical protein LOC196541	unbekannt	4
13q		LOC728215	NM_001080396	106618579	107317084	3	hypothetical protein LOC728215	unbekannt	4
13q		MBNL2	NM_144778	96672574	96844375	9	muscleblind-like 2 isoform 1	unbekannt	4
13q		MBNL2	NM_207304	96672574	96844375	7	muscleblind-like 2 isoform 3	unbekannt	4
13q		MCF2L	NM_024979	112704068	112792497	27	MCF-2 cell line derived transforming	Guaninnukleotid-Austauschfaktor	3
13q		MYO16	NM_015011	108046500	108658356	35	myosin heavy chain Myr 8	Myosin	3
13q		MYO16	NM_015011	108243864	108506406	22	myosin heavy chain Myr 8	Myosin	3
13q		PCCA	NM_000282	99533337	99980689	24	propionyl-Coenzyme A carboxylase, alpha	Untereinheit der mitochondrialen Propionyl-CoA Carboxylase	3
13q		PCID2	NM_018386	112879926	112910984	15	PCJ domain containing 2	unbekannt	4
13q		PHGDHL1	NM_177967	98651163	98836682	7	phosphoglycerate dehydrogenase like 1	unbekannt	4
13q		PROZ	NM_003891	112860968	112874695	8	protein Z, vitamin K-dependent plasma	mögl. Rolle in Hämostase	3

13q	RAB20	NM_017817	109973413	110012072	2	RAB20, member RAS oncogene family	Apikale Endocytose/Recycling	3
13q	RANBP5	NM_002271	97403929	97474551	29	RAN binding protein 5	nukleozytoplasmatischer Transport	3
13q	RAP2A	NM_021033	96884476	96918245	2	RAP2A, member of RAS oncogene family	unbekannt	4
13q	RASA3	NM_007368	113765295	113916197	24	RAS p21 protein activator 3	Aktivierung RAS p21 GTPase, Negative Regulierung RAS-Funktion	3
13q	RNF113B	NM_178861	97626039	97627522	2	ring (zinc) finger protein 113B	unbekannt	4
13q	SLC10A2	NM_000452	102494350	102517197	6	solute carrier family 10 (sodium/bile acid)	Rolle in Reabsorption von Gallensäuren, Cholesterollowerstand	3
13q	SLC15A1	NM_005073	98134056	98202909	23	solute carrier family 15 (oligopeptide)	Oligopeptidtransporter	3
13q	SOX1	NM_005986	111769913	111774021	1	SRY (sex determining region Y)-box 1	Transkriptionsfaktor	3
13q	STK24	NM_001032296	97900455	98027397	11	serine/threonine kinase 24 (STE20 homolog)	Signaltransduktionskinase	4
13q	STK24	NM_003576	97900455	97972342	11	serine/threonine kinase 24 (STE20 homolog)	Signaltransduktionskinase	4
13q	TTFDP1*	NM_007111	113287056	113343506	12	transcription factor Dp-1	Transkriptionsfaktor, Vermittlung von Proliferation und Apoptose, Amplifikation/Überexpression in Brustkrebs	2
13q	TM9SF2	NM_004800	98951728	99013277	17	transmembrane 9 superfamily member 2	unbekannt	4
13q	TMCO3	NM_017905	113193310	113252508	13	transmembrane and coiled-coil domains 3	unbekannt	4
13q	TMTCA	NM_001079669	100054090	100125104	18	transmembrane and tetratricopeptide repeat	unbekannt	4
13q	TMTCA	NM_032813	100054090	100125104	19	transmembrane and tetratricopeptide repeat	unbekannt	4
13q	TNFSF13B	NM_006573	107719977	107757366	6	tumor necrosis factor (ligand) superfamily	Zytokin, B-Zell Aktivierung	3
13q	TPP2	NM_003291	102047373	102129522	29	tripeptidyl peptidase II	Protein der proteolytischen 26S Proteasom-Kaskade	3
13q	TUBGCP3	NM_006322	112187328	112290482	22	spindle pole body protein	Protein des Spindelpolkörpers	3
13q	UPF3A	NM_023011	114065179	114089385	10	UPF3 regulator of nonsense transcripts homolog A	Teil eines post-Spleißing Komplexes mit Rolle in mRNA Kernexport und Kontrolle	3
13q	UPF3A	NM_080687	114065179	114089385	9	UPF3 regulator of nonsense transcripts homolog A	Teil eines post-Spleißing Komplexes mit Rolle in mRNA Kernexport und Kontrolle	3
13q	VGCNL1	NM_052867	100504130	100866814	44	voltage gated channel like 1	unbekannt	4
13q	ZIC2	NM_007129	99432319	99437020	3	zinc finger protein of the cerebellum 2	Transkriptionsfaktor, Rolle in u.a. zerebellarer Entwicklung	3
13q	ZIC5	NM_033132	99413275	99422179	2	zinc finger protein of the cerebellum 5	Transkriptionsfaktor, Rolle in Entwicklung	3

Tabelle 10: Gene in der Kandidatenregion 19p13.2-q12 (10556621-35934942)

Kandidaten region	GenSymbol	RefSeq- Nummer	Transkription Start	Transkription Ende	Exone	Name (englisch)	kurze Funktionsbeschreibung	Rang
19	ABHD8	NM_024527	17263940	17275234	5	abhydrolase domain containing 8	unbekannt	4
19	ABHD9	NM_024794	15198729	15204231	7	abhydrolase domain containing 9	unbekannt	4
19	ACP5	NM_001611	11546476	11550800	7	tartrate resistant acid phosphatase 5 precursor	Katalyse der Umwandlung von Orthophospho-Monoester zu Alkohol und Orthophosphat	3
19	AKAP8	NM_005858	15325334	15351603	14	A-kinase anchor protein 8	Zellzyklus-abh. Interaktion mit Proteinkinase A	3
19	AKAP8L	NM_014371	15351858	15390833	14	A kinase (PRKA) anchor protein 8-like	mögl. Rolle in Initiationsphase der DNA-Replikation	4
19	ANKRD25	NM_015493	11135945	11167496	12	ankyrin repeat domain 25	unbekannt	4
19	ANKRD41	NM_152363	17253680	17259454	10	ankyrin repeat domain 41	unbekannt	4
19	APIM1	NM_032493	16169730	16207151	12	adaptor-related protein complex 1, mu 1 subunit	Komponente eines Golgi-Komplexes mit Rolle in Endozytose und Golgiprocessing	3
19	ARMC6	NM_033415	19005470	19029980	8	armadillo repeat containing 6	unbekannt	4
19	ARRDC2	NM_001025604	17972943	17985911	8	arrestin domain containing 2 isoform 2	unbekannt	4
19	ARRDC2	NM_015683	17979976	17985911	8	arrestin domain containing 2 isoform 1	unbekannt	4
19	ASF1B	NM_018154	14091320	14108440	4	anti-silencing function 1B	Histonchaperon, mögl. Rolle in Modulation der Nukleosomenstruktur des Chromatin	3
19	ASNA1	NM_004317	12709305	12720137	7	arsA arsenite transporter, ATP-binding, homolog	Aniontransport ATPase	3
19	ATP13A1	NM_020410	19617009	19633925	26	ATPase type 13A1	unbekannt	4
19	B3GNT3	NM_014256	17766918	17785385	3	UDP-GlcNAc:betaGal	Transferase für Oligosaccharidsynthese, L-Selektin Ligand Biosynthese, Lymphozytenfunktion	3
19	BEST2	NM_017682	12724406	12730271	9	vitelliform macular dystrophy 2-like 1	Chloridkanal	3
19	BRD4	NM_014299	15218846	15252262	12	bromodomain-containing protein 4 isoform short	Steuerung chromosomaler Dynamik während Mitose	3
19	BRD4	NM_058243	15209300	15252262	20	bromodomain-containing protein 4 isoform long	Steuerung chromosomaler Dynamik während Mitose	3
19	BST2	NM_004335	17374754	17377384	5	bone marrow stromal cell antigen 2	mögl. Rolle in B-Zell Wachstum und rheumatoider Arthritis	4
19	BTBD14B	NM_052876	13090108	13112959	6	transcriptional repressor NAC1	Transkriptionsrepressor, onkogenetisches Potenzial	4
19	C19orf12	NM_001031726	34881632	34897803	3	hypothetical protein LOC83636 isoform 1	unbekannt	4
19	C19orf12	NM_031448	34881632	34898292	3	hypothetical protein LOC83636 isoform 2	unbekannt	4
19	C19orf2	NM_003796	35125264	35198453	11	RPB5-mediating protein isoform a	Negative Modulierung der Transkription durch Bindung an RNA Polymerase II	3
19	C19orf2	NM_134447	35125264	35198453	10	RPB5-mediating protein isoform b	Negative Modulierung der Transkription durch Bindung an RNA Polymerase II	3
19	C19orf39	NM_175871	11346382	11348627	2	hypothetical protein LOC126074	unbekannt	4
19	C19orf42	NM_024104	16617958	16631968	5	hypothetical protein LOC79086	unbekannt	4
19	C19orf43	NM_024038	12702453	12706529	3	hypothetical protein MGC2803	unbekannt	4
19	C19orf44	NM_032207	16468204	16493163	9	hypothetical protein LOC84167	unbekannt	4
19	C19orf50	NM_024069	18529603	18541186	5	hypothetical protein LOC79036	unbekannt	4
19	C19orf52	NM_138358	10900423	10902000	2	hypothetical protein LOC90580	unbekannt	4
19	C19orf53	NM_014047	13746256	13750586	3	hypothetical protein LOC28974	unbekannt	4
19	C19orf56	NM_016145	12639886	12641449	4	hypothetical protein LOC51398	unbekannt	4
19	C19orf57	NM_024323	13854167	13877909	8	hypothetical protein LOC79173	unbekannt	4
19	C19orf58	NM_024050	17281353	17292099	5	hypothetical protein LOC79016	unbekannt	4
19	C19orf60	NM_017967	18560607	18564146	3	hypothetical protein LOC55049	unbekannt	4
19	CACNA1A	NM_000068	13179114	13478317	48	calcium channel, alpha 1A subunit isoform 1	Spannungsabh. Calciumkanal	3
19	CACNA1A	NM_023035	13179114	13478317	48	calcium channel, alpha 1A subunit isoform 2	Spannungsabh. Calciumkanal	3



19	CALR	NM_004343	12910422	12916303	9	calreticulin precursor	Calciumbindendes Chaperon im ER, Calretikulin/Calnexin Zyklus, mögl. Rolle in Transkriptionsregulation	3
19	CALR3	NM_145046	1645086	16468003	9	calreticulin 3	Calciumbindendes Chaperon im ER, Calretikulin/Calnexin Zyklus	3
19	CARM1	NM_199141	10843252	10894448	16	coactivator-associated arginine	Methylierung von Zielproteinen, Rolle in Transkriptionsregulation, mRNA-Stabilität	3
19	CASP14	NM_012114	15024014	15027900	6	caspace 14 precursor	Caspase mit Rolle in terminaler Differenzierung von Keratinozyten, downstream Caspase der Apoptose	3
19	CC2D1A	NM_017721	13878051	13902692	29	coiled-coil and C2 domain containing 1A	Transkriptionsfaktor mit mögl. Rolle in neuronalen Zellen	4
19	CCDC105	NM_173482	14982555	14995081	7	coiled-coil domain containing 105	unbekannt	4
19	CCDC124	NM_138442	17906904	17915783	5	coiled-coil domain containing 124	unbekannt	4
19	CCDC130	NM_030818	13719752	13735106	10	coiled-coil domain containing 130	unbekannt	4
19	CCNE1	NM_001238	34994740	35007059	12	cyclin E1 isoform 1	Zyclin für CDK2, dessen Aktivität für G1/S Zellzyklustransition verantwortlich. Überexpression resultiert in chromosomaler Instabilität	2
19	CCNE1	NM_057182	34995400	35007059	10	cyclin E1 isoform 2	Zyclin für CDK2, dessen Aktivität für G1/S Zellzyklustransition verantwortlich. Überexpression resultiert in chromosomaler Instabilität	2
19	CD97	NM_001025160	14353212	14380535	19	CD97 antigen isoform 3 precursor	Rezeptor der Leukozytenaktivierung, -migration	3
19	CD97	NM_001784	14353212	14380535	18	CD97 antigen isoform 2 precursor	Rezeptor der Leukozytenaktivierung, -migration	3
19	CD97	NM_078481	14353212	14380535	20	CD97 antigen isoform 1 precursor	Rezeptor der Leukozytenaktivierung, -migration	3
19	CHERP	NM_006387	16489699	16514263	17	calcium homeostasis endoplasmic reticulum	unbekannt	4
19	CHB3	NM_054113	16133178	16145286	6	DNA-dependent protein kinase catalytic	unbekannt	4
19	CLIP2	NM_153221	19510114	19517903	8	cartilage intermediate layer protein 2	mögl. Rolle in Knorpelgest	4
19	CNN1	NM_001299	11510578	11522138	7	calponin 1, basic, smooth muscle	Regulation und Modulation der Muskelkontraktion	3
19	COMP	NM_000095	18754582	18763114	19	cartilage oligomeric matrix protein precursor	EZ-Matrix Protein	3
19	COPE	NM_007263	18871322	18891199	10	epsilon subunit of coatomer protein complex	Untereinheit des Coatomerkomplexes	3
19	COPE	NM_199442	18871322	18891199	9	epsilon subunit of coatomer protein complex	Untereinheit des Coatomerkomplexes	3
19	COPE	NM_199444	18871322	18891199	9	epsilon subunit of coatomer protein complex	Untereinheit des Coatomerkomplexes	3
19	CPAMD8	NM_015692	16864761	16998625	42	C3 and P2P-like, alpha-2-macroglobulin domain	unbekannt	4
19	CRLF1	NM_004750	18565046	18578660	10	cytokine receptor-like factor 1	Untereinheit Zytokinrezeptor, mögl. Rolle in Immunsystem und fetaler Entwicklung	4
19	CRSP7	NM_004831	16546717	16600015	3	cofactor required for Sp1 transcriptional	Transkriptionaler Koaktivator	3
19	CRTC1	NM_015321	18655424	18749758	14	mucoepidermoid carcinoma translocated 1 isoform	Transkriptionaler Koaktivator	4
19	CRTC1	NM_025021	18655424	18749758	14	mucoepidermoid carcinoma translocated 1 isoform	Transkriptionaler Koaktivator	4
19	CYP4F11	NM_021187	15884180	15906326	12	cytochrome P450, family 4, subfamily F	Monoxygenase, Reaktionskatalyse	3
19	CYP4F12	NM_023944	15644917	15668984	13	cytochrome P450, family 4, subfamily F	Monoxygenase, Reaktionskatalyse	3
19	CYP4F2	NM_001082	15849833	15869884	13	cytochrome P450, family 4, subfamily F	Monoxygenase, Reaktionskatalyse	3
19	CYP4F3	NM_000896	15612706	15632570	13	cytochrome P450, family 4, subfamily F	Monoxygenase, Reaktionskatalyse	3
19	CYP4F8	NM_007253	15587028	15601447	14	cytochrome P450, family 4, subfamily F	Monoxygenase, Reaktionskatalyse	3
19	DAND5	NM_152654	12941323	12946567	2	dante precursor	Antagonist des BMP (bone morphogenic protein)	3
19	DDX39*	NM_005804	14380630	14391171	11	DEAD (Asp-Glu-Ala-Asp) box polypeptide 39	putative RNA-Helikase, Rolle in pre-mRNA Spleißen, mRNA Kernexport	3
19	DDX49	NM_019070	18891493	18900436	13	DEAD (Asp-Glu-Ala-Asp) box polypeptide 49	unbekannt	4
19	DHPS	NM_001930	12647533	12653677	9	deoxyhypusine synthase isoform a	Enzym der Hypusin Biosynthese	3
19	DHPS	NM_013406	12647533	12653677	8	deoxyhypusine synthase isoform b	Enzym der Hypusin Biosynthese	3
19	DHPS	NM_013407	12647533	12653677	8	deoxyhypusine synthase isoform c	Enzym der Hypusin Biosynthese	3
19	DNAJB1	NM_006145	14486551	14490201	3	DnaJ (Hsp40) homolog, subfamily B, member 1	Interaktion mit HSP70	4

19	DNASE2	NM_001375	12847024	12853335	6	deoxyribonuclease II, lysosomal precursor	Hydrolyse von DNA	3
19	DNM2	NM_001005360	10689754	10803579	21	dynamitin 2 isoform 1	Mikrotubuli-assoziiertes Motorprotein	3
19	DNM2	NM_001005361	10689754	10803579	21	dynamitin 2 isoform 2	Mikrotubuli-assoziiertes Motorprotein	3
19	DNM2	NM_001005362	10689754	10803579	20	dynamitin 2 isoform 4	Mikrotubuli-assoziiertes Motorprotein	3
19	DNM2	NM_004945	10689754	10803579	20	dynamitin 2 isoform 3	Mikrotubuli-assoziiertes Motorprotein	3
19	DOCK6	NM_020812	11170972	11234128	48	dedicator of cytokinesis 6	mögl. Guaninukleotid-Austauschfaktor	4
19	ECSIT	NM_016581	11477743	11500930	8	evolutionarily conserved signaling intermediate	Signaltransduktionsprotein	4
19	EDG4	NM_004720	19595476	19600017	3	endothelial differentiation, lysophosphatidic acid receptor 3	LPA-Rezeptor, Rolle in Ca2+-Mobilisierung und PLC- $\beta$ Signalweg	3
19	ELAVL3	NM_001420	11423142	11452803	7	ELAV-like protein 3 isoform 1	Neuronenspezifisches RNA-bindendes Protein	3
19	ELAVL3	NM_032281	11423142	11452803	7	ELAV-like protein 3 isoform 2	Neuronenspezifisches RNA-bindendes Protein	3
19	ELL	NM_006532	18414472	18439318	12	elongation factor RNA polymerase II	Elongationsfaktor	3
19	ELOF1	NM_032377	11524857	11531051	4	elongation factor 1 homolog (ELF1)	Elongationsfaktor	3
19	EMR2	NM_013447	14704204	14750353	21	egf-like module containing, mucin-like, hormone	unbekannt	4
19	EMR2	NM_152916	14704204	14750353	20	egf-like module containing, mucin-like, hormone	unbekannt	4
19	EMR2	NM_152917	14704204	14750353	19	egf-like module containing, mucin-like, hormone	unbekannt	4
19	EMR2	NM_152918	14704204	14750353	18	egf-like module containing, mucin-like, hormone	unbekannt	4
19	EMR2	NM_152919	14704204	14750353	21	egf-like module containing, mucin-like, hormone	unbekannt	4
19	EMR2	NM_152920	14704204	14750353	20	egf-like module containing, mucin-like, hormone	unbekannt	4
19	EMR2	NM_152921	14704204	14750353	19	egf-like module containing, mucin-like, hormone	unbekannt	4
19	EMR3	NM_032571	14591050	14646810	16	egf-like module-containing mucin-like receptor 3	mögl. Rolle in der myeloischen Interaktion	4
19	EMR3	NM_152939	14591050	14646810	15	egf-like module-containing mucin-like receptor 3	mögl. Rolle in der myeloischen Interaktion	4
19	EPOR	NM_000121	11349474	11356019	8	erythropoietin receptor precursor	Rezeptor für EPO	3
19	EPS15L1	NM_021235	16333407	16443762	23	epidermal growth factor receptor pathway	Komponente der Clathrin-coated Pits	4
19	F2RL3	NM_003950	16860825	16863830	2	coagulation factor II (thrombin) receptor-like 3	Rezeptor für aktiviertes Thrombin oder Trypsin, mögl. Rolle in Plättchenaktivierung	3
19	FAM125A	NM_138401	17391911	17397140	9	hypothetical protein LOC93343	unbekannt	4
19	FAM129C	NM_173544	17495109	17525454	16	B-cell novel protein 1	unbekannt	4
19	FAM32A	NM_014077	16157234	16163857	4	hypothetical protein LOC26017	unbekannt	4
19	FARSA	NM_004461	12894292	12905522	13	phenylalanyl-tRNA synthetase, alpha subunit	Enzym der tRNA Synthese	3
19	FBXW9	NM_032301	12680729	12688455	10	F-box and WD-40 domain protein 9	Komponente des E3 Ubiquitinligase Komplexes	3
19	FCHO1	NM_015122	17723335	17760365	29	FCH domain only 1	unbekannt	4
19	FKBP8	NM_012181	18503567	18511538	7	FK506-binding protein 8	mögl. Rolle in den Neuronen, assoziiert mit Gedächtnisfunktion und in Regulierung der Apoptose	4
19	FKBP8	NM_012181	18503567	18515383	9	FK506-binding protein 8	mögl. Rolle in den Neuronen, assoziiert mit Gedächtnisfunktion und in Regulierung der Apoptose	4
19	FKSG24	NM_032683	18165039	18168550	5	hypothetical protein LOC84769	unbekannt	4
19	FLJ25328	NM_152483	16005563	16013945	5	hypothetical protein LOC148231	unbekannt	4
19	FLJ39501	NM_173483	15480336	15524127	14	cytochrome P450, family 2, subfamily E	unbekannt	4
19	FLJ44894	NM_001039884	20367762	20399602	4	similar to zinc finger protein 91	unbekannt	4
19	FLJ44968	NM_198537	19500719	19509393	7	hypothetical protein LOC374887	unbekannt	4
19	FLJ45910	NM_207390	14564895	14582951	12	hypothetical protein LOC388512	unbekannt	4
19	GADD45GIP1	NM_052850	12925971	12929050	2	growth arrest and DNA-damage-inducible, gamma	Negativer Regulator der G1/S Zellzyklusprogression durch Inhibierung von CDKs, möglicher Tumorsuppressor in Krebszellen	2

19	GATAD2A	NM_017660	19430396	19478789	12	GATA zinc finger domain containing 2A	Transkriptionsrepressor	4
19	GCDH	NM_000159	12862973	12871783	12	glutaryl-Coenzyme A dehydrogenase isoform a	Enzym im Aminosäure-Metabolismus	3
19	GCDH	NM_013976	12862973	12871783	12	glutaryl-Coenzyme A dehydrogenase isoform b	Enzym im Aminosäure-Metabolismus	3
19	GDF15	NM_004864	18387984	18380986	2	growth differentiation factor 15	Morphogenetisches Knochenprotein	4
19	GIPC1	NM_005716	14449570	14467944	9	regulator of G-protein signalling 19 interacting	Gewebedifferenzierung und -erhalt	4
19	GIPC1	NM_202467	14449570	14467944	8	regulator of G-protein signalling 19 interacting	mögl. Rolle in G-Protein gekoppelter Signalleitung	4
19	GIPC1	NM_202468	14449570	14467944	8	regulator of G-protein signalling 19 interacting	mögl. Rolle in G-Protein gekoppelter Signalleitung	4
19	GIPC1	NM_202469	14449570	14467944	7	regulator of G-protein signalling 19 interacting	mögl. Rolle in G-Protein gekoppelter Signalleitung	4
19	GIPC1	NM_202470	14449570	14467944	7	regulator of G-protein signalling 19 interacting	mögl. Rolle in G-Protein gekoppelter Signalleitung	4
19	GIPC1	NM_202494	14449570	14467944	6	regulator of G-protein signalling 19 interacting	mögl. Rolle in G-Protein gekoppelter Signalleitung	4
19	GLT25D1	NM_024656	17527510	17554965	12	glycosyltransferase 25 domain containing 1	unbekannt	4
19	GFSN2	NM_138501	14501381	14537792	13	glycoprotein, synaptic 2	unbekannt	4
19	GMIP	NM_016573	19601284	19615455	21	GEM interacting protein	Interaktionsprotein der Signaltransduktions	3
19	GTPBP3	NM_032620	17309363	17314530	10	GTP binding protein 3 (mitochondrial) isoform V	Modifikation der mitochondrialen tRNA	3
19	GTPBP3	NM_133644	17309363	17313874	8	GTP binding protein 3 (mitochondrial) isoform	Modifikation der mitochondrialen tRNA	3
19	HAPLN4	NM_023002	19227470	19234651	5	brain link protein 2	mögl. Rolle in Formation der EZ-Matrix	4
19	HOMER3	NM_004838	18901011	18912983	10	Homer, neuronal immediate early gene, 3	Regulation der Glutamaterezeptorfunktion	3
19	HOOK2	NM_013312	12734816	12747337	23	hook homolog 2	mögl. Rolle als Zytoskeletlinkerprotein	4
19	HSH2D	NM_032855	16104021	16130377	9	hematopoietic SH2 domain containing	T-Zell Aktivierungswegway	3
19	HSPC142	NM_001033549	17239231	17251162	9	hypothetical protein LOC29086	unbekannt	4
19	HSPC142	NM_014173	17239231	17251162	9	hypothetical protein LOC29086	unbekannt	4
19	HSPFP36	NM_001080493	11693077	11710731	4	ZFP-36 for a zinc finger protein	mögl. Rolle in Transkriptionsregulation	4
19	IER2	NM_004907	13122281	13126718	2	immediate early response 2	unbekannt	4
19	IFT30	NM_006332	18145578	18149927	7	interferon, gamma-inducible protein 30	Lysosomale Thiolreduktase	3
19	IL12RB1	NM_005535	18031370	18058742	17	interleukin 12 receptor, beta 1 isoform 1	Rezeptor für IL 12	3
19	IL12RB1	NM_153701	18042371	18058742	10	interleukin 12 receptor, beta 1 isoform 2	Rezeptor für IL 12	3
19	IL2TRA	NM_004843	14003261	14025026	14	class I cytokine receptor	Rezeptor für IL 27	3
19	ILF3*	NM_004516	10625987	10657443	18	interleukin enhancer binding factor 3 isoform b	mögl. Rolle in Translationsregulation und Transkriptionsregulation	4
19	ILF3*	NM_012218	10625987	10664095	20	interleukin enhancer binding factor 3 isoform a	mögl. Rolle in Translationsregulation und Transkriptionsregulation	4
19	ILF3*	NM_153464	10625987	10657443	18	interleukin enhancer binding factor 3 isoform c	mögl. Rolle in Translationsregulation und Transkriptionsregulation	4
19	ILVBL	NM_006844	15086784	15097577	16	ilvB (bacterial acetolactate synthase)-like	Mögl. Enzym der Aminosäure-Synthese	4
19	ILVBL	NM_176826	15086784	15096968	5	ilvB (bacterial acetolactate synthase)-like	Mögl. Enzym der Aminosäure-Synthese	4
19	INSL3	NM_005543	17788321	17793320	2	insulin-like 3 precursor	mögl. Rolle in Entwicklung des Urogenitalsystems	3
19	ISYNA1	NM_016368	18406624	18409943	11	myo-inositol 1-phosphate synthase A1	unbekannt	4
19	JAK3	NM_000215	17797960	17819800	24	Janus kinase 3	Kinase der Signaltransduktion	3
19	JAK3	NM_000215	17801596	17819800	23	Janus kinase 3	Kinase der Signaltransduktion	3
19	JUNB	NM_002229	12763309	12765125	1	jun B proto-oncogene	Transkriptionsfaktor der Wachstumsfaktorantwort, mögliche Rolle in der Pathogenese von Brustkrebs	2
19	JUND	NM_005354	18251569	18253432	1	jun D proto-oncogene	Transkriptionsfaktor, Schutz der Zellen vor p53-vermittelter Zellalterung und Apoptose	3
19	KCNN1	NM_002248	17923110	17970930	11	potassium intermediate/small conductance	Spannungs-unabh. Kaliumkanal	3
19	KIAA0892	NM_015329	19316782	19330563	8	hypothetical protein LOC23383	unbekannt	4
19	KIAA1683	NM_025249	18228907	18246235	4	KIAA1683	unbekannt	4
19	KLF1	NM_005633	12856236	12859017	3	Kruppel-like factor 1 (erythroid)	Transkriptionsfaktor	4
19	KLF2	NM_016270	16296650	16299339	3	Kruppel-like factor	Transkriptionsfaktor	4
19	KLHL26	NM_018316	18608837	18642302	3	kelch-like 26	unbekannt	4
19	LASS1	NM_021267	18840360	18867953	8	longevity assurance gene 1 isoform 1	Morphogenetisches Knochenprotein	4
19	LASS1	NM_198207	18849717	18867953	6	longevity assurance gene 1 isoform 2	Morphogenetisches Knochenprotein	4
19	LDLR	NM_000527	11061131	11105490	18	low density lipoprotein receptor precursor	LDL-Rezeptor	3

19	LOC126075	NM_001080503	11318333	11327053	14	hypothetical protein LOC126075	unbekannt	4
19	LOC126536	NM_001034966	15994437	15999723	5	hypothetical protein LOC126536	unbekannt	4
19	LOC388524	NM_001005472	23737655	23802759	4	hypothetical protein LOC388524	unbekannt	4
19	LOC401898	NM_001013691	11611590	11685382	10	hypothetical protein LOC401898	unbekannt	4
19	LOC55908	NM_018687	11209125	11213619	6	hepatocellular carcinoma-associated gene TD26	unbekannt	4
19	LOC729747	NM_001080404	12015619	12028127	5	similar to zinc finger protein 709	unbekannt	4
19	LOC90379	NM_138353	13924318	13933256	13	hypothetical protein LOC90379	unbekannt	4
19	LPHN1	NM_001008701	14119548	14177997	25	latrophilin 1 isoform 1 precursor	Rezeptor für alpha-Latrotoxin, mögl. Rolle in Exocytose	3
19	LPHN1	NM_014921	14119548	14177997	24	latrophilin 1 isoform 2 precursor	Rezeptor für alpha-Latrotoxin, mögl. Rolle in Exocytose	3
19	LPR2	NM_022737	11327106	11337372	10	lipid phosphate phosphatase-related protein type	unbekannt	4
19	LRR25	NM_145256	18362953	18369415	2	leucine rich repeat containing 25	mögl. Rolle in Immunzellaktivierung	4
19	LSM4	NM_012321	18278716	18294958	5	U6 snRNA-associated Sm-like protein 4	Interaktion mit snRNA, Rolle in pre-mRNA Spleißen	3
19	LYL1	NM_005583	13070846	13074681	4	lymphoblastic leukemia derived sequence 1	unbekannt	4
19	MAN2B1	NM_000528	12618321	12638828	24	mannosidase, alpha, class 2B, member 1	Enzym des Proteinstoffwechsels	3
19	MAP1S	NM_018174	17691302	17706324	7	BPY2 interacting protein 1	unbekannt	4
19	MAST1	NM_014975	12810347	12846765	26	microtubule associated serine/threonine kinase	mögl. Verlinkung von Zytoskelettfilamenten	4
19	MDS032	NM_018467	17187167	17191638	8	uncharacterized hematopoietic stem/progenitor	SNARE mit mögl. Rolle im Vesikeltransport	4
19	MEF2B	NM_005919	19117377	19163933	12	MADS box transcription enhancer factor 2	Transkriptionsfaktor	4
19	MGC20983	NM_145045	11392271	11406980	13	hypothetical protein LOC115948	unbekannt	4
19	MGC3207	NM_001031727	13736336	13746096	6	hypothetical protein LOC84245 isoform 1	unbekannt	4
19	MGC3207	NM_032285	13736336	13746096	6	hypothetical protein LOC84245 isoform 2	unbekannt	4
19	MORG1	NM_032332	12641568	12647646	9	mitogen-activated protein kinase organizer 1	Protein verschiedener Komplexe, Rolle im ERK Pathway, Reaktion auf Hypoxie, pre-mRNA Spleißen	4
19	MRPL34	NM_023937	17277476	17278652	2	mitochondrial ribosomal protein L34	Ribosomales Protein der Mitochondrien der 39S Untereinheit	3
19	MYO9B	NM_004145	17073469	17185104	39	myosin IXB	Myosin Motorprotein	4
19	NANOS3	NM_001098622.1	13849062	13852317	2	NANOS3 protein	Rolle in Keimzellerhalt	3
19	NCAN	NM_004386	19183781	19224061	15	chondroitin sulfate proteoglycan 3	Proteoglycan mit Rolle in Zelladhäsion und Migration	3
19	NDUFB7	NM_004146	14537889	14543886	3	NADH dehydrogenase (ubiquinone) 1 beta	Enzym des Komplexes 1 der mitochondrialen Atmungskette	3
19	NDUFA13	NM_015965	19487595	19500007	5	cell death-regulatory protein GRIM19	Enzym des Komplexes 1 der mitochondrialen Atmungskette, Rolle in Apoptose	3
19	NFIX	NM_002501	12967583	13070610	10	nuclear factor I/X (CCAAT-binding transcription factor)	Transkriptions-/Replikationsaktivierung	3
19	NOTCH3	NM_000435	15131443	15172792	33	Notch homolog 3	Rolle in intrazellulären Signalwegen, Transkriptionsaktivierung	3
19	NR2F6	NM_005234	17203693	17217151	4	nuclear receptor subfamily 2, group F, member 6	unbekannt	4
19	NWD1	NM_001007525	16691786	16789763	21	NACHT and WD repeat domain containing 1	unbekannt	4
19	NY-SAR-48	NM_001011699	17021570	17047343	11	sarcoma antigen NY-SAR-48 isoform b	unbekannt	4
19	NY-SAR-48	NM_033417	17021570	17047343	11	sarcoma antigen NY-SAR-48 isoform a	unbekannt	4
19	OCEL1	NM_024578	17198054	17201028	6	occludin/ELL domain containing 1	unbekannt	4
19	OR10H1	NM_013940	15779840	15799847	1	olfactory receptor, family 10, subfamily H	Olfaktorischer Rezeptor	3
19	OR10H2	NM_013939	15699853	15700801	1	olfactory receptor, family 10, subfamily H	Olfaktorischer Rezeptor	3
19	OR10H3	NM_013938	15713202	15714153	1	olfactory receptor, family 10, subfamily H	Olfaktorischer Rezeptor	3
19	OR10H4	NM_001004465	15920817	15921768	1	olfactory receptor, family 10, subfamily H	Olfaktorischer Rezeptor	3

19	OR10H5	NM_001004466	15765858	15766806	1	olfactory receptor, family 10, subfamily H	Olfaktorischer Rezeptor	3
19	OR1I1	NM_001004713	15058876	15059944	1	olfactory receptor, family 1, subfamily I	Olfaktorischer Rezeptor	3
19	OR7A10	NM_001005190	14812759	14813689	1	olfactory receptor, family 7, subfamily A	Olfaktorischer Rezeptor	3
19	OR7A17	NM_030901	14852237	14853167	1	olfactory receptor, family 7, subfamily A	Olfaktorischer Rezeptor	3
19	OR7A17	NM_017506	14798138	14800276	1	olfactory receptor, family 7, subfamily A	Olfaktorischer Rezeptor	3
19	OR7C1	NM_198944	14770985	14771948	1	olfactory receptor, family 7, subfamily C	Olfaktorischer Rezeptor	3
19	OR7C2	NM_012377	14913300	14914260	1	olfactory receptor, family 7, subfamily C	Olfaktorischer Rezeptor	3
19	PBX4	NM_025245	19533521	19590439	8	pre-B-cell leukemia homeobox 4	mögl. Rolle als Transkriptionsfaktor	4
19	PDE4C	NM_000923	18182347	18198285	15	phosphodiesterase 4C, cAMP-specific	Nukleotidmetabolismus	3
19	PDE4C	NM_000923	18182347	18205210	15	phosphodiesterase 4C, cAMP-specific	Nukleotidmetabolismus	3
19	PDE4C	NM_000923	18182347	18220010	16	phosphodiesterase 4C, cAMP-specific	Nukleotidmetabolismus	3
19	PDE4C	NM_000923	18182347	18227229	19	phosphodiesterase 4C, cAMP-specific	Nukleotidmetabolismus	3
19	PGLS	NM_012088	17483431	17493097	5	6-phosphogluconolactonase	Enzym Glucosetrofwechsel	3
19	PGLYRP2	NM_052890	15440462	15451312	4	peptidoglycan recognition protein L precursor	Teil des Immunostoffwechsels, Erkennung von Peptidoglykanen	3
19	PGPEP1	NM_017712	18312407	18341763	5	pyroglutamyl-peptidase I	Enzym des Aminosäure-Stoffwechsels	3
19	PK3R2	NM_005027	18125015	18142343	16	phosphoinositide-3-kinase, regulatory subunit 2	Interaktion mit Protein-Tyrosinkinasen	3
19	PKN1	NM_002741	14405165	14443679	22	protein kinase N1 isoform 2	Rolle in intrazellulären Signalwegen, mögl. Funktion im Apoptose-Signalweg	2
19	PKN1	NM_213560	14412085	14443679	22	protein kinase N1 isoform 1	Rolle in intrazellulären Signalwegen, mögl. Funktion im Apoptose-Signalweg	2
19	PLEKHF1	NM_024310	34847802	34858204	2	apoptosis-inducing protein D	Funktion im Apoptose-Signalweg	2
19	PLXAP	NM_031310	17323263	17349137	6	plasma membrane vesicle associated protein	Induzierung von Apoptose durch lysosomal-mitochondrialen Pathway	2
19	PODNL1	NM_024825	13904165	13910289	8	podocan-like 1	Rolle in Diaphragmaformation der Caveolen	3
19	POP4	NM_006627	34789040	34798547	7	POP4 (processing of precursor, S. cerevisiae)	unbekannt	4
19	PRDX2	NM_005809	12768633	12773694	6	peroxiredoxin 2 isoform a	Komponente der Ribonuklease P, die reife tRNA generiert	3
19	PRDX2	NM_181738	12771672	12773694	3	peroxiredoxin 2 isoform c	Rolle in zellulärer Redoxregulation, mögl. proliferativer Effekt und Rolle in Krebsentstehung	2
19	PRKACA	NM_002730	14063506	14089559	10	cAMP-dependent protein kinase catalytic subunit	Rolle in zellulärer Redoxregulation, mögl. proliferativer Effekt und Rolle in Krebsentstehung	2
19	PRKACA	NM_207518	14063506	14085992	10	cAMP-dependent protein kinase catalytic subunit	Phosphorylierung von Zielproteinen	4
19	PRKOSH	NM_001001329	11407268	11422782	18	protein kinase C substrate 80K-H isoform 2	Untereinheit der Glucosidase II im ER	3
19	PRKOSH	NM_002743	11407268	11422782	18	protein kinase C substrate 80K-H isoform 1	Untereinheit der Glucosidase II im ER	3
19	PTGER1	NM_000955	14444277	14447174	3	prostaglandin E receptor 1, subtype EP1	Rezeptor für Prostaglandin E2	3
19	QTRT1	NM_031209	10673130	10685040	10	queine tRNA-ribosyltransferase 1 (tRNA-guanine)	Enzym der tRNA Synthese	3
19	RAB3A	NM_002866	18168610	18175839	5	RAB3A, member RAS oncogene family	mögl. Rolle in Regulation der Exozytose	4
19	RAB3D	NM_004283	11296092	11311321	5	RAB3D, member RAS oncogene family	mögl. Rolle in Regulation der Exozytose	4
19	RAB8A	NM_005370	16083489	16106045	8	mel transforming oncogene	mögl. Rolle in vesiculärem Trafficking, Neurotransmitterausschüttung	4
19	RAD23A	NM_005053	12917653	12925455	9	UV excision repair protein RAD23 homolog A	Rolle in Postreplikations-Reparatur UV-geschädigter DNA, Basenexzisionsreparatur, mögliche Rolle im Ubiquitin vermittelten Proteolyse Pathway	2
19	RFX1	NM_002918	13933351	13978097	21	regulatory factor X1	Regulatorischer Transkriptionsfaktor für u.a. MHCII Expression	3
19	RFXANK	NM_003721	19164007	19173678	10	regulatory factor X-associated	Regulatorischer Transkriptionsfaktor für u.a. MHCII Expression	3
19	RFXANK	NM_134440	19164007	19173678	9	regulatory factor X-associated	Regulatorischer Transkriptionsfaktor für u.a. MHCII Expression	3
19	RGL3	NM_001035223	11366000	11391004	19	ral guanine nucleotide dissociation	unbekannt	4

19	RLN3	NM_080864	14000016	14002783	2	relaxin 3 preproprotein	Hormon	3
19	RNA5EH2A	NM_006397	12778427	12785462	8	ribonuclease H2, large subunit	Untereinheit RNase H II, die spezifisch RNA von RNA-DNA-Hybriden degradiert	3
19	RPL18A	NM_000980	17831726	17835124	5	ribosomal protein L18a	Ribosomales Protein der 60S Untereinheit	3
19	RTBDN	NM_001080997	12797295	12802207	6	retbindin isoform 1	unbekannt	4
19	RTBDN	NM_031429	12797295	12807230	7	retbindin isoform 2	unbekannt	4
19	SAMD1	NM_138352	14060228	14062232	6	sterile alpha motif domain containing 1	unbekannt	4
19	SF4	NM_172231	19248321	19292307	14	splicing factor 4	Rolle in pre-mRNA Spleißen	3
19	SFRS14	NM_001017392	18964176	19005368	11	splicing factor, arginine/serine-rich 14	mögl. Rolle in pre-mRNA Spleißen	4
19	SFRS14	NM_014884	18962701	19005368	11	splicing factor, arginine/serine-rich 14	mögl. Rolle in pre-mRNA Spleißen	4
19	SIN3B	NM_015260	16801217	16852164	20	SIN3 homolog B, transcription regulator	Transkriptionsrepressor	4
19	SIC1A6	NM_005071	14921900	14944729	9	solute carrier family 1 (high affinity)	Aminosäure-Transporter	3
19	SIC2B4A2	NM_178526	19035807	19084697	8	solute carrier family 25, member 42	unbekannt	4
19	SIC27A1	NM_198580	17442299	17477977	12	solute carrier family 27, member 1	Lipidmetabolismus	3
19	SIC35E1	NM_024881	16523581	16544193	6	solute carrier family 35, member E1	Putativer Transporter	4
19	SIC44A2	NM_020428	10597228	10616232	22	CTL2 protein	Putativer Transporter	4
19	SIC5A5	NM_000453	17843781	17865897	15	solute carrier family 5 (sodium iodide)	Natrium/Iodid Transporter	3
19	SMARCA4	NM_003072	10932605	11083953	35	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	Komponente des Chromatin-Remodeling Komplex SNF/SWI, Transkriptionsaktivierung, mögl. Rolle in Regulation der Zellproliferation, mögl. Tumorsuppressor, kann BRCA1 binden	2
19	SPC24	NM_182513	11117169	11127458	5	spindle pole body component 24 homolog	Komponente eines Kinetochor-assoziierten Komplexes mit Rolle in Chromosomen-segregation und Spindelkontrollpunkt-Aktivität	4
19	SSBP4	NM_001009998	18391135	18406459	17	single stranded DNA binding protein 4 isoform b	unbekannt	4
19	SSBP4	NM_032627	18391135	18406459	18	single stranded DNA binding protein 4 isoform a	unbekannt	4
19	STPX10	NM_003765	13116223	13121987	8	syntaxin 10	SNARE, Vesikeltransport	3
19	SYCE2	NM_001105578.1	12870893	12890151	5	Synaptonemal complex central element protein 2	Komponente des zentralen Elements des synaptonemalen Komplexes	3
19	SYDE1	NM_033025	15079213	15086789	8	synapse defective 1, Rho GTPase, homolog 1	unbekannt	4
19	TM6SF2	NM_001001524	19236173	19245074	10	transmembrane 6 superfamily member 2	unbekannt	4
19	TMED1	NM_006858	10804113	10807983	4	interleukin 1 receptor-like 1 ligand precursor	unbekannt	4
19	TMEM161A	NM_017814	19091439	19110267	12	hypothetical protein LOC54929	unbekannt	4
19	TMEM16H	NM_020959	17295033	17306638	18	hypothetical protein LOC57719	unbekannt	4
19	TMEM38A	NM_024074	16632937	16660816	6	transmembrane protein 38A	unbekannt	4
19	TMEM59L	NM_012109	18584711	18592822	8	brain-specific membrane-anchored protein	mögl. Rolle in Hirnfunktion	4
19	TNPO2	NM_013433	12671014	12694078	25	transportin 2 (importin 3, karyopherin beta 2b)	mögl. Rolle im nukleären Proteinimport	4
19	TPM4	NM_003290	16048324	16074408	8	tropomyosin 4	Regulation der Muskelkontraktion, Stabilisierung des Aktin-Zytoskeletts	3
19	TRA16	NM_176880	19173223	19175238	5	TR4 orphan receptor associated protein TRA16	mögl. Transkriptionsrepressor	4
19	TRMT1	NM_017722	13076714	13088332	16	TRM1 tRNA methyltransferase	Enzym des tRNA-Stoffwechsels	3
19	TSPAN16	NM_012466	11267823	11298672	7	transmembrane 4 superfamily member 16	Signaltransduktionsprotein	4
19	TSSK6	NM_032037	19486027	19487469	1	serine/threonine protein kinase SSTK	Rolle in Spermienproduktion/-reifung, DNA-Kondensation	3
19	TXX1L6	NM_138454	17427233	17432725	2	thioredoxin-like 6	unbekannt	4
19	UBA52	NM_001033930	18543613	18549270	5	ubiquitin and ribosomal protein L40 precursor	Proteinmodifizierung mit folgender Degradation bzw. Chromatinstrukturhalt, Regulation der Genexpression, Stressreaktion und DNA-Reparatur	2

19	UBA52	NM_003333	18543613	18549270	5	ubiquitin and ribosomal protein L40 precursor	Proteinmodifizierung mit folgender Degradation bzw. Chromatinstruktur, Regulation der Genexpression, Stressreaktion und DNA-Reparatur	2
19	UNC13A	NM_001080421	17573136	17680401	44	unc-13 homolog A	Rolle in synaptischen Vesikel-Priming/-Reifung und Neurotransmitterausschüttung	3
19	UNQ501	NM_198536	11314457	11317722		hypothetical protein LOC374882	unbekannt	4
19	UPF1	NM_002911	18803743	18840039	24	regulator of nonsense transcripts 1	Komponente eines post-Spleifing Komplexes	3
19	UPF1	NM_002911	18803743	18840039	24	regulator of nonsense transcripts 1	Komponente eines post-Spleifing Komplexes	3
19	UQCRCFS1	NM_006003	34390006	34395954	2	ubiquinol-cytochrome c reductase, Rieske	Komplex III der mitochondrialen Atmungskette	3
19	USHBP1	NM_031941	17221848	17236544	13	Usher syndrome 1C binding protein 1	unbekannt	4
19	YIPF2	NM_024029	10894443	10900357	10	Yip1 domain family, member 2	unbekannt	4
19	ZNF100	NM_173531	21697407	21742163	5	zinc finger protein 100	mögl. Rolle in Transkriptionsregulation	4
19	ZNF101	NM_033204	19640662	19652363	4	zinc finger protein 101	mögl. Rolle in Transkriptionsregulation	4
19	ZNF136	NM_003437	12134918	12161064	4	zinc finger protein 136 (clone pHZ-20)	mögl. Rolle in Transkriptionsregulation	4
19	ZNF14	NM_021030	19682280	19704921	4	zinc finger protein 14	mögl. Rolle in Transkriptionsregulation	4
19	ZNF20	NM_021143	12103802	12121440	4	zinc finger protein 20 (KOX 13)	mögl. Rolle in Transkriptionsregulation	4
19	ZNF208	NM_007153	21940736	21985555	7	zinc finger protein 208	mögl. Rolle in Transkriptionsregulation	4
19	ZNF253	NM_021047	19837713	19855293	4	zinc finger protein 253	mögl. Rolle in Transkriptionsregulation	4
19	ZNF254	NM_004876	24061815	24104494	4	zinc finger protein 254 isoform 2	mögl. Rolle in Transkriptionsregulation	4
19	ZNF257	NM_033468	22027188	22064563	5	zinc finger protein 257	mögl. Rolle in Transkriptionsregulation	4
19	ZNF333	NM_032433	14661869	14692772	12	zinc finger protein 333	mögl. Rolle in Transkriptionsregulation	4
19	ZNF429	NM_001001415	21480276	21512919	4	zinc finger protein 429	mögl. Rolle in Transkriptionsregulation	4
19	ZNF43*	NM_003423	21779591	21810810	4	zinc finger protein 43 (HTF6)	mögl. Rolle in Transkriptionsregulation	4
19	ZNF430	NM_025189	20995336	21033493	5	zinc finger protein 430	mögl. Rolle in Transkriptionsregulation	4
19	ZNF431	NM_133473	21116727	21180645	5	zinc finger protein 431	mögl. Rolle in Transkriptionsregulation	4
19	ZNF433	NM_001080411	21986531	21987525	4	zinc finger protein 433	mögl. Rolle in Transkriptionsregulation	4
19	ZNF439	NM_152262	11837843	11841306	3	zinc finger protein 439	mögl. Rolle in Transkriptionsregulation	4
19	ZNF44	NM_016264	12219006	12266637	8	zinc finger protein 44	mögl. Rolle in Transkriptionsregulation	4
19	ZNF440	NM_152357.2	11786106	11807016	4	Zinc finger protein 440	mögl. Rolle in Transkriptionsregulation	4
19	ZNF441	NM_152355	11738906	11754301	5	zinc finger protein 441	mögl. Rolle in Transkriptionsregulation	4
19	ZNF442	NM_030824	12321184	12337475	6	zinc finger protein 442	mögl. Rolle in Transkriptionsregulation	4
19	ZNF443	NM_005815	12401519	12412819	4	zinc finger protein 443	mögl. Rolle in Transkriptionsregulation	4
19	ZNF486	NM_052852.2	20139022	20169984	4	Zinc finger protein 486	mögl. Rolle in Transkriptionsregulation	4
19	ZNF490	NM_020714	12547644	12582595	5	zinc finger protein 490	mögl. Rolle in Transkriptionsregulation	4
19	ZNF491	NM_152356	11770390	11780306	3	zinc finger protein 491	mögl. Rolle in Transkriptionsregulation	4
19	ZNF492	NM_020855.1	22609078	22642312	4	Zinc finger protein 492 (Fragment)	mögl. Rolle in Transkriptionsregulation	4
19	ZNF493	NM_001076678	21371770	21402136	4	zinc finger protein 493 isoform 3	mögl. Rolle in Transkriptionsregulation	4
19	ZNF493	NM_175910	21371770	21402136	2	zinc finger protein 493 isoform 1	mögl. Rolle in Transkriptionsregulation	4
19	ZNF506	NM_001099269.1	19765642	19793558	3	Zinc finger protein 506	mögl. Rolle in Transkriptionsregulation	4
19	ZNF536	NM_014717	35555167	35740805	5	zinc finger protein 536	mögl. Rolle in Transkriptionsregulation	4
19	ZNF563	NM_145276	12289303	12305502	4	zinc finger protein 563	mögl. Rolle in Transkriptionsregulation	4
19	ZNF564	NM_14976	12497539	12523317	4	zinc finger protein 564	mögl. Rolle in Transkriptionsregulation	4
19	ZNF625	NM_145233	12116708	12128529	4	zinc finger protein 625	mögl. Rolle in Transkriptionsregulation	4
19	ZNF626	NM_001076675	20594584	20636242	4	zinc finger protein 626 isoform 1	mögl. Rolle in Transkriptionsregulation	4
19	ZNF626	NM_145297	20619348	20636242	4	zinc finger protein 626 isoform 2	mögl. Rolle in Transkriptionsregulation	4
19	ZNF627	NM_145295	11569326	11590974	4	zinc finger protein 627	mögl. Rolle in Transkriptionsregulation	4
19	ZNF653	NM_138783	11455245	11477654	9	zinc finger protein 653	mögl. Rolle in Transkriptionsregulation	4
19	ZNF675	NM_138330	22627547	236611857	4	zinc finger protein 675	mögl. Rolle in Transkriptionsregulation	4
19	ZNF676	NM_001001411	22153742	22171593	3	zinc finger protein 676	mögl. Rolle in Transkriptionsregulation	4
19	ZNF681	NM_138286	23717999	23733479	3	zinc finger protein 681	mögl. Rolle in Transkriptionsregulation	4
19	ZNF682	NM_001077349	19976226	20011064	4	zinc finger protein 682 isoform 2	mögl. Rolle in Transkriptionsregulation	4
19	ZNF682	NM_033196	19976226	20011277	4	zinc finger protein 682 isoform 1	mögl. Rolle in Transkriptionsregulation	4
19	ZNF69	NM_021915	11859669	11886132	5	zinc finger protein 69	mögl. Rolle in Transkriptionsregulation	4
19	ZNF700	NM_144566	11896899	11922578	4	zinc finger protein 700	mögl. Rolle in Transkriptionsregulation	4

19	ZNF708	NM_021269	21265802	21304052	4	zinc finger protein 15-like 1 (KOX 8)	mögl. Rolle in Transkriptionsregulation	4
19	ZNF709	NM_152601	12435742	12456632	4	zinc finger protein 709	mögl. Rolle in Transkriptionsregulation	4
19	ZNF714	NM_182515	21056848	21099723	7	zinc finger protein 714	mögl. Rolle in Transkriptionsregulation	4
19	ZNF763	NM_001012753	11936868	11952198	4	zinc finger protein 440 like	mögl. Rolle in Transkriptionsregulation	4
19	ZNF791	NM_153358	12582752	12601676	4	zinc finger protein 791	mögl. Rolle in Transkriptionsregulation	4
19	ZNF799	NM_001080821.1	12361829	12372794	4	Zinc finger protein 799	mögl. Rolle in Transkriptionsregulation	4
19	ZNF85	NM_003429	20897933	20925343	4	zinc finger protein 85 (HPP4, HTF1)	mögl. Rolle in Transkriptionsregulation	4
19	ZNF91	NM_003430	2332337	23370109	4	zinc finger protein 91	mögl. Rolle in Transkriptionsregulation	4
19	ZNF93	NM_031218	19872786	19906766	4	zinc finger protein 93	mögl. Rolle in Transkriptionsregulation	4
19	ZNF99	NM_001080409	22730846	22744624	8	zinc finger protein 99	mögl. Rolle in Transkriptionsregulation	4
19	ZSWIM4	NM_023072	13767273	13804044	13	zinc finger, SWIM domain containing 4	unbekannt	4



## 9 ERKLÄRUNGEN

### 9.1 Ehrenwörtliche Erklärung

„ Ich, Monika Huber, erkläre hiermit, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet. Insbesondere habe ich nicht die entgeltliche Hilfe von Vermittlungs- bzw. Beratungsdiensten (Promotionsberater oder andere Personen) in Anspruch genommen. Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeit erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen. Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt“.

---

Ort, Datum

Monika Huber

### 9.2 Erklärung zur Anregung und Überwachung der Dissertation

„ Ich, Monika Huber, erkläre hiermit, dass Herr Prof. Dr. Bernhard Weber die vorliegende Dissertation angeregt und überwacht hat.“

---

Ort, Datum

Monika Huber

## 10 DANKSAGUNG

An dieser Stelle möchte ich allen herzlich danken, die durch ihre Unterstützung und ihr Interesse zur Vollendung dieser Arbeit beigetragen haben.

Meinem Doktorvater, Prof. Dr. Bernhard Weber, danke ich für die Themenstellung, die kompetente Anleitung sowie die vielen Gespräche und Anregungen, die zum Gelingen dieser Arbeit beigetragen haben. Durch seine hervorragende Betreuung war es mir möglich, Einblicke in die Forschung und das wissenschaftliche Arbeiten zu erhalten.

Mein besonderer Dank gilt Dr. Lars Fritsche für die große Unterstützung bei theoretischen Fragestellungen und Hilfe mit diversen Computerprogrammen. Seine ständige Bereitschaft zu Gesprächen hat zur Weiterentwicklung meines Themas beigetragen und die Arbeit in der vorliegenden Form entstehen lassen.

Weiterhin danke ich Daniela Wagner für die Einarbeitung in die Labortechniken und die vielen technisch-praktischen Hilfestellungen zu jeder Zeit meiner Arbeit. Ebenso gilt mein Dank allen weiteren Mitarbeiterinnen und Mitarbeitern des Instituts für Humangenetik für die freundliche Aufnahme in ihr Team, die große Hilfsbereitschaft bei praktischen Fragen und Umsetzungen sowie für das angenehme Arbeitsklima.

Für die Durchsicht und Unterstützung bei der Formatierung dieser Arbeit danke ich meiner Schwester Marion, Thomas Maag und Günter Heinrichs.

Danke an Markus für die vielen aufbauenden Worte und die großartige Motivation; dass er immer an mich glaubte und mir einfach wundervoll zur Seite stand.

Abschließend möchte ich mich von Herzen bei meinen Eltern bedanken, die mich auf meinem Lebensweg bis hierher stets in jeder Hinsicht liebevoll unterstützt und begleitet haben.